How to beat diffusion: explorations of energetics and spatial relationships in microbial ecosystems.

Thesis by

Grayson Lee Chadwick

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

(Defended 06/09/2020)

Grayson Lee Chadwick
ORCID: 0000-0003-0700-9350

ACKNOWLEDGEMENTS

I would like to thank current and past members of the Orphan lab who have helped make my time in the lab so enjoyable: Abigail Green-Saxena, Jeffrey Marlow, David Case, Elizabeth Trembath-Reichert, John and Paul Magyar, Kat Dawson, Ally Pasulka, Lizzy Wilbanks, Roland Hatzenpichler, Silvan Scheller, Derek Smith, Josh Steele, Jennifer Glass, Sean Mullin, Kyle Metcalfe, Sujung Lim, Aditi Narayanan, Yongzhao Guo, Ranjani Murali, Haley Sapers, Daan Speth, Fabai Wu, Antoine Cremiere, Alon Philosof and Yamini Jangir. A special thanks to Anne Dekas, Shawn McGlynn and Hank Yu, whose interactions in the Orphan lab have been the most meaningful personally and influential scientifically. Also, a particularly important thanks to Yunbin Guan and Stephanie Connon—without their expertise and professionalism none of this work would have been possible. An additional thanks to Fernanda Jiménez Otero; our collaboration on a number of projects has led to some of the most exciting and satisfying results of my scientific career, and none of it would have been possible without her amazing work.

I would also like to thank the professors who were instrumental in my early time at Caltech and set me on my current path: Dave Stevenson, Tom Tombrello, Pamela Bjorkman and Rob Phillips. Thanks also to Justin Bois, who taught me most of the useful things I can do with a computer, and whose contributions to education at Caltech is immeasurable. Visiting professors to the department have also greatly enriched my time at Caltech, including Martin Ackermann, Andreas Kappler, Michi Wagner, Jeff Gralnick, Steve Zinder, Hiroyuki Imachi, Scott Dawson, Margaret McFall-Ngai, Ned Ruby and Shana Goffredi.

I have been incredibly fortunate to have the continued support of truly wonderful committee members: Jared Leadbetter, Dianne Newman, Woody Fischer and Victoria Orphan. All four have helped shape the way I think about science in their own ways. Their inclusion of me as a teaching assistant in the Woods Hole Microbial Diversity course and the Agouron Geobiology course have been a very special part of my time as a graduate student. As my primary advisor, Victoria trusted me with a wide array of different projects, allowing me the freedom to explore beyond the normal topics studied in the lab, and this freedom has been incredibly important for my growth and a scientist. I'm forever grateful for that trust.

Finally, an all-important thanks to my parents Maxine and Oliver, my sister Dana, and Usha. Their loving support through this time was the most important of all.

ABSTRACT

This thesis investigates four microbial systems, with a particular focus for how spatial considerations shape the behavior and evolution of microorganisms. After a general introduction in Chapter 1, Chapter 2 presents the results of experiments demonstrating how cellular activity varies through space within an anode-reducing biofilm. Chapter 3 presents a comprehensive comparative genomic analysis of all known marine anaerobic methanotrophic archaea, supporting the notion that these organisms share many energetic similarities with the anode reducing organisms in Chapter 2. These are interpreted as specific adaptations to life in highly structured microbial communities. Chapter 4 describes the enrichment and characterization of a new member of the purple sulfur bacteria, and the adaptations that may improve substrate acquisition beyond the normal limitations of diffusion. Chapter 5 describes the convergent evolution of novel Complex I gene clusters that have incorporated new proton pumping subunits, and the modifications made to the protein structure to facilitate the incorporation of these new subunits into the quaternary structure of the complex.

"A rival of mine once complained that my stories begin awkwardly and end untidily. I am willing to admit to many faults, but I will not burden my conscience with that one. All my tales are true, drawn from life, and a life story is not a tidy thing. It is a half-tamed horse that you seize on the run and ride with knees and teeth clenched, and then you regretfully slip off as gently and safely as you can, always wondering if you could have gone a few metres more."

-Karen Lord, *Redemption in Indigo*

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Chadwick, G. L., Otero, F. J., et al. (2019). "NanoSIMS imaging reveals metabolic stratification within current-producing biofilms". In: *Proceedings of the National Academy of Sciences*, *116*(41), 20716-20724. doi:10.1073/pnas.1912498116

G. L. C participated in conception of the project, study design, carried out labwork, analysis and writing of the manuscript.

2. Chadwick, G. L., et al. (2018). "Convergent evolution of unusual complex I homologs with increased proton pumping capacity: energetic and ecological implications". In: *The ISME journal*, *12*(11), 2668-2680. doi: 10.1038/s41396-018-0210-1

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

3. He, X., Chadwick, G. L., et al. (2020a). "Controls on interspecies electron transport and size limitation of anaerobically methane oxidizing microbial consortia" Submitted.

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

4. He, X., Chadwick, G. L., et al. (2020b). "Spatially Resolved Electron Transport through Anode-Respiring Geobacter sulfurreducens Biofilms: Controls and Constraints". Submitted.

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

5. Bird, L. R., Dawson, K. S., Chadwick, G. L., et al. (2019) "Carbon isotopic heterogeneity of coenzyme F430 and membrane lipids in methane-oxidizing archaea". In: *Geobiology* 17.6: 611-627. doi: 10.1111/gbi.12354

G. L. C participated in the writing of the manuscript.

6. Bublitz, D. C., Chadwick, G. L., et al. (2019). "Peptidoglycan production by an insect-bacterial mosaic." *Cell* 179.3: 703-712. doi: 10.1016/j.cell.2019.08.054

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

7. Boyd, J. A., Jungbluth, S. P., Leu, A. O., Evans, P. N., Woodcroft, B. J., Chadwick, G. L., et al. (2019). Divergent methyl-coenzyme M reductase genes in a deep-subseafloor Archaeoglobi. *The ISME journal*, 13(5), 1269-1279. doi: 10.1038/s41396-018-0343-2

G. L. C participated in the analysis and writing of the manuscript.

8. He, X., Chadwick, G., et al. (2019). Microbial interactions in the anaerobic oxidation of methane: model simulations constrained by process rates and activity patterns. *Environmental microbiology*, 21(2), 631-647. doi: 10.1111/1462-2920.14507

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

9. Dekas, A. E., Fike, D. A., Chadwick, G. L., et al. (2018). Widespread nitrogen fixation in sediments from diverse deep-sea sites of elevated carbon loading. *Environmental microbiology*, 20(12), 4281-4296. doi: 10.1111/1462-2920.14342

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

10. McGlynn, S. E., Chadwick, G. L., et al. (2018). Subgroup characteristics of marine methane-oxidizing ANME-2 archaea and their syntrophic partners as revealed by integrated multimodal analytical microscopy. *Appl. Environ. Microbiol.*, *84*(11), e00399-18. doi: 10.1128/AEM.00399-18

G. L. C participated in conception of the project, data acquisition, analysis and writing of the manuscript.

11. Dekas, A. E., Connon, S. A., Chadwick, G. L., et al. (2016). Activity and interactions of methane seep microorganisms assessed by parallel transcription and FISH-NanoSIMS analyses. *The ISME journal*, *10*(3), 678-692. doi: 10.1038/ismej.2015.145

G. L. C participated in data acquisition, analysis and writing of the manuscript.

12. Scheller, S., Yu, H., Chadwick, G. L., et al. (2016). Artificial electron acceptors decouple archaeal methane oxidation from sulfate reduction. *Science*, *351*(6274), 703-707. doi: 10.1126/science.aad7154

G. L. C participated in data acquisition, analysis and writing of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS AND/OR TABLES

Chapter 1

1. INTRODUCTION

Ecology is concerned with the interactions of organisms with each other and their environment. It follows that the physical position of an individual relative to other organisms and the abiotic factors that affect its growth is a crucial component in understanding its ecology. In microbial ecology, the distances over which important interactions occur are often small—nanometers to centimeters—with niches defined by steep gradients of nutrients or light that can change drastically over these spatial scales.

A major focus of this thesis is the advancement of our ecological understanding of the organisms responsible for the anaerobic oxidation of methane (AOM) in marine sediments. This process is carried out by a symbiotic partnership between an anaerobic methane-oxidizing (ANME) archaea and a sulfate reducing bacteria (SRB). The ANME archaea oxidize methane to CO_2, producing 8 electrons, that are transferred to the SRB for sulfate reduction. These organisms form dense spherical biofilms, or aggregates, affording them a spatial proximity that facilitates their coupled metabolisms.

ANME-SRB aggregates occur throughout the world's oceans, particularly along continental margins, where methane produced at depth rises up through diffusion or advection and meets sulfate-rich ocean water. These opposing gradients of methane and sulfate define the sulfate-methane transition zone (SMTZ), a narrow horizon in marine sediments where methane is oxidized with the reduction of sulfate. Here we see the effect of space in microbial ecology on multiple scales. First, the supply of methane and sulfate, a result of diffusion and/or advection, determines the position of the SMTZ in which the

ANME-SRB symbiosis can thrive. Second, the transfer of methane-derived electrons between partners necessitates the close physical association of cells on the scale of microns.

This metabolic partnership is a common form of microbial interaction. Cross-feeding, or syntrophy, where the end product of one organism's metabolism is the starting point of another's, is commonplace in anaerobic microbiology. Often times these interactions are obligate for both parties—the first organism is inhibited by a buildup of its waste if a second is not there to consume it, and the second organism has no growth substrate without the activity of the first. The transport of metabolites from one organism to another occurs mainly by diffusion, requiring steep gradients over short distances for rapid transport. Therefore, the positioning of organisms in structured microbial communities really matters, and developing strategies to mix themselves—so that no producer is ever far from its nearest consuming partner—is thought to be an important aspect of their evolution (Schink and Thauer, 1988).

The expectation that a well-mixed community optimizes metabolic cross-feeding is at odds with the apparent spatial organization of many ANME-SRB aggregates, which can be highly segregated with a dense ANME population at the center surrounded by a shell of SRB. ANME cells at the center of these aggregates are many cell distances from their nearest syntrophic partner, raising questions about their ability to remain active without a nearby sink for their metabolic waste. Reaction-diffusion modeling implemented with electron transfer based on diffusible chemical compounds such as H_2, formate, or acetate suggested that this arrangement could not support the growth of most cells in the consortia (Alperin and Hoehler, 2009; Orcutt and Meile, 2008), only cells at the interface between the two populations. This was seen as a major limitation for the growth of these organisms, and these shell-type aggregates were interpreted as a terminal stage of growth of these communities (Schink and Stams, 2006).

The work presented in this thesis began with the development of techniques to probe metabolic activity in ANME-SRB aggregates on the single-cell level, allowing us to test the hypothesis that cells in these consortia experience a growth decrease with distance from their nearest partner (McGlynn et al., 2015). By incubating sediment samples from methane seeps with ^{15}N labeled ammonium, we were able to observe the uptake of ^{15}N, a proxy for a cell's anabolic activity. Contrary to our expectations, we saw no significant decrease in cellular activity with increasing distance to syntrophic partners, as all the theory and modelling based on diffusive electron transfer had predicted.

A reanalysis of the few publicly available genomes revealed that ANME encode large multi-heme cytochromes (MHCs). MHCs are often utilized by organisms that carry out direct electron transfer, where electrons obtained from oxidizing their carbon source are shuttled through the MHCs to external electron acceptors such as insoluble metal oxides or the anode in an electrochemical cell. A model organism that carries out this metabolism, *Geobacter sulfurreducens*, can form thick biofilms on suitably poised anodes. Direct electron transfer using MHCs and other conductive biomolecules was thought to enable the growth of *G. sulfurreducens* cells at the surface of these biofilms, tens of microns away from their electron acceptor (Reguera et al., 2006; Sun et al., 2016), distances much greater than could be supported by electron transfer via diffusible intermediates. Based on this knowledge from *G. sulfurreducens*, our finding of large MHC proteins encoded in ANME genomes, and the apparent insensitivity of ANME or SRB to similar distances from their partners, we proposed that the ANME-SRB syntrophy was carried out by direct electron transfer through MHCs, alleviating the spatial constraints thought to be imposed by diffusive electron transport (McGlynn et al., 2015).

This argument was predicated on the assumption that *G. sulfurreducens* biofilms do, in fact, support the growth of cells many cell lengths away from their terminal electron acceptor. While this idea was pervasive in the literature (Reguera et al., 2006; Sun et al., 2016), the metabolic activity of cells in these biofilms had not been quantitatively assessed. In Chapter 2, we present the results of experiments carried out in collaboration with researchers at the University of Minnesota to investigate this question, using the techniques we developed to investigate cellular activity in ANME-SRB aggregates. The results of these experiments were surprising, with *G. sulfurreducens* biofilms repeatedly demonstrating a clear decrease in cellular activity with increasing distance from the anode. These results are important not only for our interpretation of the patterns of activity within the ANME-SRB aggregates, but also for the broader bioelectrochemistry community that is interested in the factors that limit current generation in microbial fuel cells.

Reaction-diffusion modeling has contributed important insights into the factors that could control cellular activity in ANME-SRB aggregates and other structured microbial communities. We have continued advancing these models, expanding on the comparison between electron transport through diffusive chemical intermediates and direct electron transport through MHCs. Using the data presented in McGlynn et al. 2015, additional modelling efforts provided support for our direct electron transport hypothesis (He et al., 2019). Subsequent single-cell ^{15}N incorporation data from exceptionally large ANME-SRB aggregates allowed us to further refine our models of electron transfer mechanisms (He et al., 2020b). Similarly, we used the *G. sulfurreducens* data that we generated in Chapter 2 to constrain new models that accurately reproduce the activity profiles observed in *G. sulfurreducens* biofilms (He et al., 2020a). A major conclusion from these studies is that direct electron transfer is not a panacea for spatial limitations—cells utilizing direct electron transfer still suffer a growth penalty with increasing distance

from their terminal electron acceptor. However, this penalty occurs over a larger distance than that from diffusion-based electron transfer, allowing direct electron transfer to support metabolic interactions across spatial scales that would not be feasible with diffusible intermediates alone.

The genomic observation of large MHC proteins in ANME was another key piece of information supporting our hypothesis that direct electron transfer was occurring in ANME-SRB consortia. However, the genomic coverage of the ANME was sparse in 2015, with some groups such as ANME-2C and -3 having no sequenced representatives at all. This poor genomic coverage raised questions about whether or not this model could be extended to all ANME groups. In Chapter 3 we combined our genomic sequencing efforts with other groups throughout the world studying these organisms to create the most complete ANME genome comparison to date, including all known subgroups. This work highlights those genomic features that set ANME apart from their closely related methanogenic relatives, as well as those that differentiate ANME groups from one another. The preponderance of bioenergetic systems expected to be used for direct electron transfer remains one of the cardinal features of the ANME genomes, and we present a comprehensive review of the energy metabolism of these organisms.

Chapter 4 focuses on a different type of microbial ecosystem, the salt marsh sediments that support the growth of sulfide-oxidizing organisms. Much like the SMTZ in methane seeps, sulfide oxidizing organisms exist in a narrow horizon where sulfide, produced at depth, diffuses up and intersects opposing gradients of chemical oxidants such as nitrate and oxygen, or light, which supports the growth of sulfide oxidizing phototrophs. The requirement for both sulfide and oxidizing power limits this niche to thin horizons within the sediment or overlying water.

Normally sulfide-oxidizing organisms are at the mercy of diffusion for the supply of their electron donors and acceptors. A remarkable exception to this is *Thiovulum majus*, an exceptionally fast swimming sulfide oxidizing chemotroph. *T. majus* is thought to beat diffusion by tethering itself in biofilms and using its flagella to pull oxygen-rich water towards itself, increasing its supply of electron acceptor. In this chapter we describe the enrichment and characterization of a novel sulfide oxidizing phototroph that has similar motility behaviors to *T. majus*. We hypothesize that this novel genus has adapted a similar strategy to increase its uptake of sulfide by facilitating the upward flow of sulfide-rich porewater. This would allow "*Candidatus* Thiorhodosphaera celerrima" to inhabit a new niche, defined by high sulfide and high light, whereas other organisms carrying out similar metabolism would have to choose one or the other due to the opposing gradients of these two resources. This proposal is further explored *in silico* by modelling approaches of varying complexity that demonstrate a significant increase in sulfide uptake.

In Chapter 5 we examine a different type of spatial scale: the nanometer-scale modifications required when protein complexes are adapted to new functions. During the genome comparison study in Chapter 3 we discovered a modified version of respiratory complex I in a variety of unrelated organisms. This complex is best known from mitochondria where it mediates electron transfer from NADH to ubiquinol, coupling this electron transfer to proton translocation, which occurs through three homologous proton pumping subunits. These complexes are broadly distributed throughout the tree of life, where they represent crucially important energy conserving steps in many diverse microbial metabolisms.

The modified complex I's we describe in this chapter were found to encode an additional copy of NuoM, one of the three proton pumping subunits. These "2M" gene clusters were found in a diverse group of

Archaea and Bacteria, and phylogenetic analysis revealed that this modification evolved convergently at least three separate times. Based on the crystal structure we were able to identify structural modifications to other subunits of the complex to physically accommodate the incorporation of a fourth site of proton translocation. The wide diversity of host organisms in which these complexes were discovered suggests that their increased proton pumping capacity might play a range of roles, important in different ways for different metabolisms.

A unifying feature of all these studies is an attempt to understand how microbial evolution has circumvented spatial limitations. In ANME-SRB consortia the limitation of cellular activity that would arise from diffusive transport of electrons is avoided by the acquisition of direct electron transfer machinery. In "*Ca.* T. celerrima" the slow delivery of sulfide by diffusion is supplemented by motility-based fluid flow. In the 2M complex I evolutionary pressure to increase proton pumping stoichiometry required physical changes in the protein structure to make way for an additional pumping subunit.

The different way these evolutionary transitions occur highlights two major modes of evolution: *de novo* invention, and horizontal transfer of functional systems. In ANME, both appear to have played a role in the adaptation to direct electron transfer. Their membrane-bound *c*-type cytochrome reductases are constructed by novel modifications to the hydrogenases of closely related methanogens, and the mechanism of electron transport through their outer surface layer appears to be unique. Therefore, both these systems were most likely invented by the ancestors of modern ANME to achieve direct electron transfer in the context of archaeal cell biology. On the other hand, a very close homolog of OmcZ, an important MHC found in *G. sulfurreducens* that is secreted into the extracellular space during growth on electrodes is found in ANME genomes. The close relation between these proteins found in organisms

from different domains of life suggests recent horizontal transfer enable this metabolism. The 2M complex I's exhibit both modes of evolution as well. These complexes have evolved independently three times, in all cases using the same physical extension mechanism. Yet, these three unrelated groups have also been distributed between divergent phyla, suggesting that while they are not prohibitively difficult to reinvent *de novo*, organisms can also benefit from acquiring them premade. In the final case of "*Ca.* T. celerrima", even though these organisms co-occur with *T, majus* in sulfidic sediments, it does not appear that the convergence of this behavior is the result of gene transfer between the two organisms.

These studies represent a holistic approach to developing an ecological understanding of microorganisms. With careful consideration of genome-encoded bioenergetic systems we can formulate coherent hypotheses that capture how environmentally providedavailable energy sources are converted through these systems into biologically useful forms of energy such as ATP and ion motive force. The explicit modeling of cellular energetics in the spatial context of their physical surroundings—be those diffusive gradients of nutrients from the environment, substrates cross-fed from a partner organism, or redox potentials imposed by an electrochemical cell—is a powerful tool for better understanding the environmental pressures that organisms have adapted to. Examining cellular activity in the context of spatial relationships allows us to ground-truth the hypotheses developed from genomic insights and modeling. Finally, comparing how these energy generating systems have changed during their vertical inheritance, or during their acquisition through horizontal transfer, provides a detailed evolutionary understanding of how modern organisms have adapted to their current niche.

Alperin MJ, Hoehler TM. 2009. Anaerobic methane oxidation by archaea/sulfate-reducing bacteria aggregates: 1. Thermodynamic and physical constraints. *American Journal of Science* **309**:869–957.

He X, Chadwick G, Kempes C, Shi Y, McGlynn S, Orphan V, Meile C. 2019. Microbial interactions in the anaerobic oxidation of methane: model simulations constrained by process rates and activity patterns. *Environmental microbiology* **21**:631–647.

He X, Chadwick GL, Jiménez-Otero F, Orphan VJ, Meile C. 2020a. Spatially Resolved Electron Transport through Anode-Respiring Geobacter sulfurreducens Biofilms: Controls and Constraints. *Submitted*.

He X, Chadwick GL, Kempes CP, Orphan VJ, Meile C. 2020b. Controls on interspecies electron transport and size limitation of anaerobically methane oxidizing microbial consortia. *Submitted*.

McGlynn SE, Chadwick GL, Kempes CP, Orphan VJ. 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* **526**:531–535.

Orcutt B, Meile C. 2008. Constraints on mechanisms and rates of anaerobic oxidation of methane by microbial consortia: process-based modeling of ANME-2 archaea and sulfate reducing bacteria interactions.

Reguera G, Nevin KP, Nicoll JS, Covalla SF, Woodard TL, Lovley DR. 2006. Biofilm and Nanowire Production Leads to Increased Current in Geobacter sulfurreducens Fuel Cells. *Appl Environ Microbiol* **72**:7345–7348. doi:10.1128/AEM.01444-06

Schink B, Stams AJM. 2006. Syntrophism among Prokaryotes In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes. New York, NY: Springer New York. pp. 309–335. doi:10.1007/0-387-30742-7_11

Schink B, Thauer RK. 1988. Energetics of syntrophic methane formation and the influence of aggregation. *Granular anaerobic sludge* 5–17.

Sun D, Chen J, Huang H, Liu W, Ye Y, Cheng S. 2016. The effect of biofilm thickness on electrochemical activity of Geobacter sulfurreducens. *International Journal of Hydrogen Energy* **41**:16523–16528. doi:10.1016/j.ijhydene.2016.04.163

Chapter 2

2. NANOSIMS IMAGING REVEALS METABOLIC STRATIFICATION WITHIN CURRENT-PRODUCING BIOFILMS

Grayson L Chadwick[1]*, Fernanda Jiménez Otero[2]*, Jeffrey A Gralnick[2,3], Daniel R Bond[2,3], and Victoria J Orphan[1]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, 91125, USA;
[2]BioTechnology Institute, [3]Department of Plant and Microbial Biology, University of Minnesota, Saint Paul, MN, 55108, USA.
*These authors contributed equally to this work.

Abstract

Metal-reducing bacteria direct electrons to their outer surfaces where insoluble metal oxides or electrodes act as terminal electron acceptors, generating electrical current from anaerobic respiration. *Geobacter sulfurreducens* is a commonly enriched electricity-producing organism, forming thick conductive biofilms that magnify total activity by supporting respiration of cells not in direct contact with electrodes. Hypotheses explaining why these biofilms fail to produce higher current densities suggest inhibition by formation of pH, nutrient, or redox potential gradients; but these explanations are often contradictory, and a lack of direct measurements of cellular growth within biofilms prevents discrimination between these models. To address this fundamental question, we measured the anabolic activity of *G. sulfurreducens* biofilms using stable isotope probing coupled to nanoscale secondary ion mass spectrometry. Our results demonstrate that the most active cells are at the anode surface, and that this activity decreases with distance, reaching a minimum 10 μm from the electrode. Cells nearest the electrode continue to grow at their maximum rate in weeks-old biofilms 80 μm thick, indicating nutrient or buffer diffusion into the biofilm is not rate-limiting. This pattern, where highest activity occurs at the electrode and declines with each cell layer, is present in thin biofilms (<5 μm), fully grown biofilms (>20 μm), and at different anode redox potentials. These results suggest a growth penalty is associated with respiring insoluble electron acceptors at micron distances, which has important implications for improving microbial electrochemical devices as well as our understanding of syntrophic associations harnessing the phenomenon of microbial conductivity.

Introduction

Microbial communities in nature exist predominantly as biofilms (Flemming et al., 2016; Stoodley et al., 2002). These dense cell aggregates coat mineral surfaces and multicellular organisms in aquatic environments, sediments, soils, and the subsurface, controlling the majority of metal, sulfur, nitrogen, and carbon cycling (Battin et al., 2003; Ehrlich, 2016; Islam et al., 2004; Tadanier et al., 2005). Due to their high cell densities, diffusion of buffering agents and respiratory substrates to inner layers is limited (de Beer et al., 1994; Domozych and Domozych, 2008; Gantzer et al., 1988; Ramsing et al., 1993; Stoodley et al., 2002, 1998). Stratification of biofilms into phenotypically (Sauer et al., 2002) and metabolically (Wilmes et al., 2009) separate regions is well documented, creating structured microhabitats that alter respiration of organic matter (Battin et al., 2003), biodegradation of substrates (Lemaire et al., 2008), and inhibition by antibiotics (Flemming et al., 2016).

Biofilms comprised of metal-reducing bacteria represent a special case, as a subset of these organisms create between-cell conductive pathways directing electrons to acceptors many microns away. Electrodes colonized by such bacteria can oxidize organic wastes or power remote devices in a host of new electrochemical technologies, but the fundamental limits to conductive electrode biofilm metabolic activity are not understood (Sleutels et al., 2012). Unlike biofilms growing on a surface, where cells near the medium interface have optimal access to nutrients, a conductive community faces a paradox; cell layers close to an electrode have best access to the electron acceptor, but those farther from the electrode have best access to soluble electron donors and buffering agents diffusing in from the medium (Bond et al., 2002). If conductivity is rate-limiting, cell metabolism should be most rapid near the electrode. If diffusion is limiting, most cell activity is hypothesized to occur near the top.

Geobacter sulfurreducens is a model organism used to study extracellular electron transfer and biofilm conductivity (Bond et al., 2002; Holmes et al., 2004; Jung and Regan, 2007; Lee et al., 2008; Tejedor Sanz et al., 2018; Tender et al., 2002), as it can be grown using solid anodes, insoluble metal oxides, or in syntrophic granules with other bacteria as terminal electron acceptors (Summers et al., 2010). Under optimal conditions with a suitably poised electrode, *G. sulfurreducens* will form conductive biofilms over 20 cell layers thick within 3 days and produce current densities on the order of 1 $mA \cdot cm^{-2}$ (Inoue et al., 2011; Marsili et al., 2008a; Stephen et al., 2014a; Sun et al., 2015). A universal observation during *G. sulfurreducens* electrode experiments is that in the early stages of colonization, each new layer of cells results in a nearly proportional increase in total electrical current, suggesting that the first few layers have sufficient access to both medium nutrients and electron disposal pathways. As biofilm thickness exceeds about 10 μm, however, the addition of new biomass fails to proportionally increase total current (Bonanni et al., 2013; Sun et al., 2016). Despite its importance in engineering better water treatment technologies and improving anaerobic digestion, the location of the active cell population, and the fundamental limitation controlling electrical current production in these communities remains unknown.

The most common method used to detect stratification within electrode biofilms is cell viability staining in conjunction with confocal microscopy. These methods, based on dyes that bind DNA in cells with intact or compromised membranes, yield conflicting results. Many studies report dead-staining cells near the anode surface, and live-staining cells at the interface with growth medium (Sun et al., 2017, 2016, 2015). However, others find the opposite, with live-staining cells at the anode and dead-staining cells distant from the electrode (Nevin et al., 2008; Schrott et al., 2014), or observe live-staining cells throughout (Franks et al., 2012a; Reguera et al., 2006; Steidl et al., 2016). A limitation of the stain-based approach is that positively charged propidium meant to indicate "dead" cells will accumulate inside live

respiring cells in response to high membrane potential (Kirchhoff and Cypionka, 2017), and DNA-based dyes bind to extracellular nucleic acids often present in biofilms, confounding attempts to resolve the central question of where metabolic activity is occurring.

Long-range electron transfer from cells to electrodes requires active metabolism coupled to a complex conductive matrix composed of pili (Reguera et al., 2006), multiheme *c*-cytochromes (Inoue et al., 2011; Stephen et al., 2014a; Wang et al., 2019), and polysaccharides (Rollefson et al., 2011), and studies have attempted to detect metabolic stratification by measuring expression of key genes within biofilms. Through physical separation of inner (0-20 μm) from outer (20-40 μm) biofilm layers followed by microarray analysis (Franks et al., 2010), a modest downregulation in genes associated with metabolic activity was detected in outer biofilm layers. In contrast, fluorescent proteins driven by promoters for cytochromes, pili, or TCA cycle genes did not reveal significant differentiation between inner and outer layers in *G. sulfurreducens* electrode biofilms imaged by confocal microscopy (Franks et al., 2012b).

While imaging and transcriptional methods have produced conflicting results, approaches focused on *c*-type cytochrome oxidation state or redox potential within *G. sulfurreducens* biofilms consistently reveal gradients that suggest increased activity near the electrode. Multiheme *c*-type cytochromes have characteristic UV/Vis and Raman spectroscopic signatures when in oxidized vs. reduced states, enabling measurement of cytochrome oxidation state using either visual light (Jain et al., 2011; Liu et al., 2011) or confocal Raman microscopy (Lebedev et al., 2014; Robuschi et al., 2017, 2013). These spectral studies demonstrate that *G. sulfurreducens* catabolic activity results in reduction of cytochromes, and reveal a higher ratio of reduced cytochromes in layers farther from the anode surface. Measurements using microelectrodes mirror Raman-based cytochrome profiles, showing that redox potential also becomes

lower with distance from the anode surface (Babauta et al., 2012). While not direct assessments of cellular activity *per se*, reduced cytochromes and lower redox potential should limit metabolic activity of cells farther from the electrode.

The opposite conclusion is proposed based on modeling of pH gradients within current-producing biofilms. Substrate oxidation coupled to electron removal at the electrode produces protons that must diffuse out of the biofilm, and many models predict pH levels should become inhibitory close to the electrode surface. Measurements with pH sensitive dyes (E. Franks et al., 2009) or microelectrodes (Babauta et al., 2012) have confirmed pH can decrease slightly from the outer surface of a thick (>50 μm) biofilm to an anode surface, where it could become inhibitory. While most models suggest diffusion of electron donors should be adequate in all regions of the biofilm (Torres et al., 2008), measurements of acetate concentrations within a very thick (>100 μm) *G. sulfurreducens* biofilm using NMR (Renslow et al., 2013) detected a gradual decrease in acetate concentration from the bulk solution to the base of the biofilm, further indicating a diffusional limitation could occur near the electrode.

Sophisticated biofilm models and indirect measurements continue to predict these contradictory hypotheses, ranging from the idea of active biomass near the anode surface becoming limited by conductivity as cells grow farther out (Kato Marcus et al., 2007), vs. a nearly opposite configuration where pH- or nutrient-dependent effects limit growth near the anode (Torres et al., 2008). The recent discovery of multiple redox potential-specific electron transfer pathways in *G. sulfurreducens* (Levar et al., 2017, 2014; Zacharoff et al., 2016) raises a third possibility that microhabitats within the biofilm support different populations of cells utilizing different metabolic machinery, allowing uniform growth throughout analogous to metabolic differentiation reported in *Shewanella* biofilms (Teal et al., 2006). This

leaves open the fundamental mystery of what restricts cell growth in conductive microbial communities that depend on extracellular electron transfer.

To conclusively identify regions of growth within electrode-reducing biofilms and test the hypothesis that metabolic stratification occurs, we directly imaged *G. sulfurreducens* anabolic activity in biofilms at sub-cellular resolution by coupling stable isotope probing with nanometer scale secondary ion mass spectrometry (nanoSIMS). Biofilms were grown on polished graphite electrodes at two different redox potentials under conditions optimized to reproducibly capture cells during exponential current increase at a very early stage (~5 μm thickness), after the onset of current plateau (~20 μm), and after 2 weeks of continuous biofilm accumulation (~80 μm) (Marsili et al., 2010, 2008b). These experimental conditions are highly reproducible and have been shown to produce biofilms expressing all the conductive components common to *G. sulfurreducens* biofilms reported in the literature (Hallberg et al., n.d.; Jiménez Otero et al., 2018). At these various points in biofilm development, standard growth medium was substituted for stable isotope enriched medium containing ^{15}N-enriched ammonium as the N source, ^{13}C-enriched acetate as C source and electron donor, and ^{2}H-enriched water serving as a passive tracer of anabolism (Kopf et al., 2016). ^{15}N, ^{13}C and ^{2}H incorporation was visualized at ~100 nm spatial resolution, allowing for assignment of activity at the scale of single cells within biofilms at various stages of development. In all cases and redox potentials, the highest anabolic activity occurred at the electrode surface, with a near-linear decrease in activity with distance from the electrode. This distance-dependent decline was also evident in biofilms only a few cells thick, as well as in 2-week old biofilms, suggesting that electron transfer over distance, rather than diffusion of nutrients from the medium, is the primary cause of current limitations in these environments.

Results

Metabolic activity decreases from the electrode surface towards the outer edge of the biofilm.

To analyze anabolic activity within current producing biofilms, *G. sulfurreducens* was grown with graphite electrodes poised at +240 mV vs. SHE as the sole terminal electron acceptor. Prior work has shown that after an initial exponential phase where both cells and current increase at a similar rate, biomass on the electrode continues to increase, but current does not, suggesting that only a subset of the biofilm is actively growing (Marsili et al., 2008b). We first labeled cells during this characteristic 'plateau' phase, when the exponential phase of current production had ceased. 83 hours into the experiment, the growth medium was replaced by growth medium containing ^{15}N-enriched ammonium, ^{13}C-enriched acetate and ^{2}H-enriched water. Current production by biofilms after addition of isotopically labeled medium paralleled those exposed to a regular medium control (**Figure 1A**).

After an incubation period corresponding to the known doubling time of *G. sulfurreducens* (6.2 ± 0.7 h, n=7), biofilms were fixed, stained and embedded in resin for imaging by electron microscopy (**Figure 1B**) and nanoSIMS analysis (**Figure 2** and **Figure 3**). Biofilms not exposed to isotopes were incubated in isotopically labeled substrates after fixation as a control for non-specific binding and abiotic incorporation.

Electron microscopy showed biofilms 15-20 μm thick, with cells in the inner layers oriented perpendicularly to the electrode surface (Fig. 1B, 89 hours of growth). Beyond ~10 μm, cells lost this orientation and appeared less tightly packed. These ultrastructural features were similar to those previously demonstrated for *G. sulfurreducens* grown on electrodes (Inoue et al., 2011; Stephen et al.,

2014b). Thin sections of resin-embedded biofilms were deposited on silicon wafers and analyzed for $^{13}C/^{12}C$, $^{15}N/^{14}N$, and $^{2}H/^{1}H$ stable isotope incorporation.

In all nanoSIMS analyses, the greatest cellular activity (isotope enrichment) was near the electrode surface, and activity approached zero at a distance >10 μm from the electrode. The $^{14}N^{12}C^-$ ion image was best for visualizing *G. sulfurreducens* biomass because of the lack of nitrogen in the embedding resin (**Figure 2A**). The heavy nitrogen ion $^{15}N^{12}C^-$, as well as the ^{15}N fractional abundance images, demonstrate high anabolic activity where cells are adjacent to the anode (**Figure 2B** and **C**). This maximum cellular activity at the anode interface decreased with distance to a minimum at 10 μm, followed by a second small peak in cellular anabolic activity close to 15 μm from the electrode. The activity associated with this second peak was more heterogeneous than what was documented near the electrode surface, but was a consistently observed feature under these incubation conditions.

The ^{15}N fractional abundance for all nanoSIMS acquisitions of three biological replicates grown at +240 mV are shown in Figure 2D. The features of these isotope enrichment curves within the ~20μm biofilm were reproducible, including an early peak in activity at the anode surface, a linear decrease with distance to a minimum isotope enrichment value, and the occurrence of a secondary small peak between this minimum and the outer surface of the biofilm.

The fractional abundances of ^{15}N, ^{13}C and ^{2}H depicted in Figure 3A-C all showed clear enrichment at the biofilm-anode interface, with ^{15}N, ^{13}C, and ^{2}H incorporation decreasing with distance from the anode surface (**Figure 3**). Although these independent isotope tracers will be incorporated in different proportions into the major classes of biomolecules (*i.e.* lipids, proteins, *etc.*), the general activity pattern

remained the same for ^{13}C, ^{2}H and ^{15}N enrichment. Here we present data from the ^{15}N enrichment as this yielded the highest signal-to-noise ratio, with little to no exogenous nitrogen in the resin to dilute the signal from the stable isotope probing experiment (Kopf et al., 2015).

With a fractional abundance of ^{15}N added to growth media at 0.06, a *G. sulfurreducens* cell (^{15}F=0.0036) doubling in 6 hours would reach a ^{15}N fractional abundance value of ~0.03 in our experiments assuming all N came from the added NH_4^+. Consistent with this prediction, peak values of fractional abundance were as high as 0.025 at the anode surface at the end of the 6 h incubation, allowing us to infer that cells closest to the electrode were doubling at nearly their maximum reported rate during the stable isotope probing experiment, even though they were positioned at the base of the biofilm farthest from the growth medium.

Exponentially growing biofilms a few cell layers thick already display anabolic activity gradients.

Early in growth experiments, when current and biomass is increasing exponentially and biofilms are only ~5 μm thick, all cells are hypothesized to be doubling at maximal rates. Samples from this growth phase were analyzed to determine anabolic activity, and test if such biofilms showed evidence of metabolic stratification. The thickness of biofilms harvested during this early phase (**Figure 4A**) ranged between 2-4 μm (**Figure 4B** and **C**, 35 hours of growth), and all exhibited a similar magnitude and trend of isotope incorporation (**Figure 4D**).

The peak enrichment value, reflecting the maximal growth rate of cells in thin biofilms, was indistinguishable from thicker biofilms. Importantly, the region of fastest growth occurred at the same location in thin biofilms as in older biofilms. This indicated that cells in the upper layers of thick biofilms

did not inhibit growth of their brethren at the anode surface, such as by consuming acetate before it diffused to the inner layers, or by acidification. Surprisingly, in these thin biofilms, a decline in metabolic activity was still detected with increasing distance, following the same trend as seen in thicker biofilms. Based on these observations, distance from the anode is a key variable slowing cell growth even at short distances, compared to the presence of additional bacteria or thickness of the biofilm. Additional nanoSIMS images of ^{15}N incorporation and plots for exponential and plateau phase can be found in Supplemental Material (**Figure S1**).

Lowering redox potential of the anode increases the gradient in anabolic activity.

Lowering the redox potential of an anode makes electron transfer by *G. sulfurreducens* less favorable. Under these conditions, the slower metabolic rate of cells could reduce demands on the conductive network or slow consumption of nutrients relative to diffusion into the biofilm, extending the zone where growth is possible. Conversely, if redox potential simply decreases with distance from the anode without a concomitant change in metabolic rate, starting at a lower potential would shrink the region where growth is possible. To directly test how cellular activity patterns were affected by anode potential, we analyzed the isotope incorporation patterns of *G. sulfurreducens* biofilms grown at -100 mV vs. SHE compared to those grown under the more optimal +240 mV. Four separate biofilms were grown at this lower potential, subjected to labeled media at various points during or after exponential current increase to create labeled biofilms ~5 μm and ~15 μm thick, and analyzed as before with isotope incorporation data binned every half micron from the surface of the anode (**Figure 5A-E**).

The major similarity between these biofilms and those grown at +240 mV was the dominant trend of metabolic activity decreasing with distance from the anode surface. However, the decline in activity

occurred much more rapidly when the anode potential was lower (-100 mV), with no initial peak in activity. Additionally, the second peak of activity beyond 10 μm was absent, yielding a much simpler pattern which decreased rapidly with distance from the anode. Overall, the use of a lower redox potential reduced the distance from the electrode where activity was detected.

Increased driving force at the electrode surface does not immediately change anabolic activity patterns.

The difference between biofilms grown at -100 mV vs. +240 mV indicated a strong impact of anode potential on metabolic stratification in biofilms. As many microbial electrochemical systems cycle or shift the anode potential during experiments or treatments, we investigated the ability of these biofilms to respond to a sudden transition in potential. Biofilms were grown at -100 mV until current plateau stage, isotopically labeled growth medium was added as before, and then the electrode potential was changed to the more favorable +240 mV. The resulting ^{15}N fractional abundance plots resembled those grown and labeled at -100 mV, including a steep decrease in isotope incorporation with distance, and lack of significant peaks in distant regions of the biofilm (**Figure 5F**). This indicated that over the short term, the more favorable potential could not be transmitted farther from the anode, or the conductive network in these biofilms could not be remodeled within 6 h to allow the increased driving force at the base of the biofilm to be used by cells more distant from the electrode as observed in biofilms initiated at +240 mV (**Figure 2**).

Extended biofilm operation does not change the location of metabolically active cells.

These results indicate that as biofilms at both high and low redox potentials add additional biomass, cells in the outer layers of the biofilm are at a disadvantage compared to their counterparts at the biofilm-anode interface. After the biofilm reaches a characteristic thickness, additional layers cannot contribute,

producing the signature of a constant current plateau. Microbial electrochemical applications often maintain biofilms for weeks in this constant-current mode (Schrott et al., 2014; Sun et al., 2017, 2016, 2015), while the data presented thus far are derived from much younger biofilms. One hypothesis that could explain earlier reports of the majority of live cells located at the top of biofilms is that during longer experiments with thicker biofilms, the activity pattern changes. To test this possibility, four replicate *G. sulfurreducens* biofilms were grown for at least two weeks with normal growth medium on electrodes poised at +240 mV vs. SHE. Unlabeled medium was exchanged as needed to keep nutrient levels in excess and prevent starvation, maintaining current production above 500 $\mu A \cdot cm^{-2}$ for this entire period. After this extended operation, biofilms were supplied with isotopically labeled medium for 6 hours and prepared as before.

Electron microscopy revealed much larger biofilms, often over 80 μm in thickness. However, unlike the younger biofilms, cells in top layers of the long-term biofilm showed evidence of lysis or consumption of cellular material, with low cellular electron density documented in all four replicates (**Figure 6A** and **Figure S2**). The frequency of intact cells increased with depth into the biofilm, and at the anode-biofilm interface, the majority of cells appeared intact, as determined by cell morphology and electron density in the cytoplasm. These patterns were consistent with the measurements of Schrott *et al.* (Schrott et al., 2014), who found the DNA content of *Geobacter* biofilms increased steadily as biofilms were grown for similar periods of time, but the RNA:DNA ratio dropped sharply, consistent with degradation of anabolic capacity.

Representative nanoSIMS isotope ratio images collected from various positions in the biofilm demonstrated that cells located near the anode were the only bacterial cells showing metabolic activity.

Despite the large amount of cell mass in the upper portion of the biofilm, the peak activity level near the electrode was similar to those measured for 5 μm and 15 μm thick biofilms (**Figure 6B**). Cells in outer layers, far from the anode, showed very little evidence of anabolic activity, with only occasional weak cellular enrichment signals. Consistent with our interpretation of cell lysis in upper layers, cells with higher electron density also contained substantially higher $^{14}N^{12}C$ counts relative to the top of the biofilm, reflecting high cellular concentrations of nitrogen containing biomolecules in cells adjacent to the electrode (**Figure 6A** and **Figure S2**). These results indicated that for consistently poised and maintained electrode biofilms, active growth continued near the electrode surface, pushing older cells out to the biofilm-medium interface, with surprisingly no apparent inhibition or contribution from nearly 80 μm of cells above the active layer.

Are cellular activity profiles due to cell growth only at the electrode surface?

Our isotope labeling experiments were conducted within an estimated doubling of *G. sulfurreducens* to minimize the potential for altering cellular activity-distance profiles due to active cell division, however it is possible that cells at the anode surface are dividing faster than the 6 hour labeling. To verify that the observed activity patterns in our experiments represent snapshots of cellular activity within communities that are spatially stable during the labeling period, we performed incubations in which biofilms were labeled for approximately two hours with 3 times the level of isotope enrichment added to the medium. In this experiment, nanoSIMS analysis of a biofilm with similar final current densities as those presented in Figures 2 and 3 produced a similar activity profile after only a 2 hour isotope pulse (~1/3 of a doubling), demonstrating that the observed pattern of isotope enrichment was not simply due to physical displacement of *Geobacter* cells from the electrode surface during biofilm growth, but rather is an activity

profile that represents cells having lower anabolic activity as distance from the electrode increases (**Figure S3**).

Can Geobacter *respire in the absence of anabolic activity?*

The incorporation of isotopically labeled substrates requires synthesis of new cellular material from N, C, or H, and it is important to consider that some cells could actively respire and contribute to current production, but not actively divide. Fast-growing heterotrophs like *Streptococcus* and *Escherichia coli* have well-documented abilities to uncouple catabolic from anabolic metabolism (Russell and Cook, 1995), but this is less well documented in bacteria during extracellular respiration. To examine the capacity of *G. sulfurreducens* to produce current in the absence of anabolic activity, we treated plateau-stage biofilms with tetracycline, an inhibitor of protein synthesis. In the presence of excess acetate and other nutrients, current production by tetracycline-treated biofilms decreased 95%. When medium was replaced to wash tetracycline from the biofilms, respiration recovered to near original levels. These experiments suggested that current production and anabolic activity were tightly coupled in these biofilms (**Figure S4**).

Discussion

The results presented here demonstrate that *G. sulfurreducens* biofilms experience a growth penalty with increasing distance from the anode surface. This result is robust at both the level of biofilm maturation and the reduction potential of the anode. A similar rate of ^{15}N incorporation is observed in cells near the anode surface during exponential phase and in thick biofilms after current plateau, indicating that cells in the upper layers of the biofilm do not significantly inhibit inner layers by consuming substrate or producing toxic byproducts. This pattern, with anabolic activity dominating a relatively narrow zone at the anode surface, was true even after weeks of biofilm growth, indicating that the dominant synthesis

of new biomass throughout all stages occurs at the anode, pushing old inactive biomass outward. This also shows that cells inside the biofilm continue to grow exponentially even as the current production rate of the biofilm as a whole reaches "stationary phase". Thus, current producing biofilms are not a monolithic and homogenous entity, but varied and heterogeneous, with metabolically stratified regions. This heterogeneity has implications for transcriptional, proteomic, and metagenomic studies that treat entire biofilms as a single community.

The transition between an exponential increase in current to a plateau phase correlates with our findings of two major physiological regions within the biofilm. First, our results show that at distances of approximately 10 μm from the anode, anabolic activity is reduced to a minimum, regardless of the potential at which the electrode is poised. Second, the regular orientation of the cells perpendicular to the anode surface breaks down at approximately this same point. Biofilms grown at high potential do continue to increase total current at a slow but steady rate (**Figure 1A**), and exhibit a previously unrecognized secondary zone of modest anabolic activity in outermost layers (**Figure 2D**), while low potential biofilms do not exhibit any further current increase within this secondary zone (**Figure 5A-D**).

Our nanoSIMS activity data agrees with redox potential and cytochrome oxidation state measurements, and studies finding significant differences between the inner ~10 μm of these biofilms vs, outer layers (Franks et al., 2010; Steidl et al., 2016; Stephen et al., 2014a). Specifically, previous results using viability or respiratory stains suggested that while cells are alive throughout the biofilm, outer layers had different metabolic rates based on transcriptional profiles (38) or changes in protein abundance (Inoue et al., 2011; Stephen et al., 2014a). These previous results are consistent with the anabolic gradient presented in

Figure 2, producing a quantitative measure of what were previously qualitative observations. Cells appear to only experience the set potential near the anode surface, while cells located farther from electrodes experience an increasingly reduced pool of extracellular cytochromes to act as their electron acceptor (Babauta et al., 2012; Jain et al., 2011; Lebedev et al., 2014; Liu et al., 2011; Robuschi et al., 2017, 2013; Snider et al., 2012), which ultimately makes electron transfer unfavorable to support growth. The low redox potential in these outer layers still appears to allow some acetate oxidation to proceed, based on the slow increase in current that occurs over subsequent days to weeks. We hypothesize that the second zone of anabolic activity in this region >10 μm from the electrode represents a shift to low potential-specific mode of growth, enabled by recently discovered systems such as the quinol oxidase CbcL that operates around potentials of -100 mV (Zacharoff et al., 2016). The fact that electrode surfaces poised at potentials near -100 mV fail to develop a distant zone of growth supports this model, as potential should more rapidly decrease to an unfavorable value in these biofilms. Alternatively, this second increase in anabolic activity could represent a switch to reliance on long-range electron transfer pathways such as those previously proposed to be dependent on the presence of cytochrome polymers (Wang et al., 2019) and/or pili (Reguera et al., 2006; Steidl et al., 2016).

These data represent the first direct quantitative measurement of *G. sulfurreducens* anabolic activity at high spatial resolution within conductive biofilms. Our results highlight the uniqueness of current-producing biofilms, which stratify in a different pattern than canonical biofilm models, with the most anabolically active cells in bottom layers where dead or persister cells are typically found (Flemming et al., 2016; Stoodley et al., 2002). The combined effect of factors that could limit cellular metabolism within electroactive biofilms favors growth at the biofilm-anode interface. Localization of the most active cells at the electrode interface dispels the notion that pH or substrate limitation strongly inhibits growth of

cells in the inner layers of these biofilm, despite models predicting such an effect (Babauta et al., 2012; E. Franks et al., 2009; Renslow et al., 2013; Torres et al., 2008). In fact, in some cases, the highest rates of growth occurred within pockets or pits inside the graphite electrodes, where diffusion to cells was most severely restricted (*i.e.* **Figure S1A**). While these data cannot resolve which components enable extracellular electron transfer, our results constrain future models to ones that account for the major mode of biofilm growth occurring in layers near the anode that push biomass outward into increasingly less favorable conditions.

These results support the widely held conclusion that *G. sulfurreducens* biofilms are conductive at distances exceeding 10 μm, but reveal that metabolism-at-a-distance comes with a cost, which is eventually inhibitory to growth. While the exact pattern of activity will vary between reactor designs, substrate concentrations, and especially buffering capacity of the media, the data presented here can be used to reconcile conflicting models of electroactive biofilms. Follow-up studies using mutants lacking conductive appendages such as polymerized OmcS cytochromes (Wang et al., 2019), or redox potential-specific oxidases such as CbcL, could test new predictions regarding the molecular basis for growth within biofilms. With the growing number of environmental Bacteria and Archaea shown to be capable of direct extracellular electron transfer, either for microbial growth on metal oxide materials, or to mediate syntrophic growth between microbes e.g. (McGlynn et al., 2015; Summers et al., 2010; Wegener et al., 2015), enhancing our mechanistic understanding of growth by current-producing microorganisms is an important goal of environmental microbiology. Our findings place new limitations on the extracellular electron transfer process and help constrain the extent to which it can support active cell growth in these increasingly important microbial systems.

Materials and Methods

Geobacter *growth*

All experiments were performed using our laboratory strain of *G. sulfurreducens* PCA, with each experiment using freshly streaked single colonies from freezer stocks. Anaerobic NB medium (0.38 g/L KCl, 0.2 g/L NH_4Cl, 0.069 g/L $NaH_2PO_4H_2O$, 0.04 g/L $CaCl_22H_2O$, 0.2 g/L $MgSO_47H_2O$, 1% v/v trace mineral mix, pH 6.8, buffered with 2 g/L $NaHCO_3$ and flushed with oxygen-free 80:20 N_2:CO_2 gas mix) with 20 mM acetate as electron donor, 40 mM fumarate as electron acceptor was used to grow liquid cultures from colony picks. Three cm^2 graphite electrodes were polished using 1500 grade sand paper (Gator Finishing Products) followed by 0.05 micron alumina slurry (Buehler), and cleaned via sonication to serve as working electrodes. Sterile three-electrode conical reactors containing 15 mL of acceptor-free medium with 40 mM acetate as the electron donor and 50 mM NaCl to equilibrate salt concentrations were flushed with a mix of N_2-CO_2 gas (80:20, v/v) until the O_2 concentration reached less than 2 ppm. Cultures were prepared by inoculating multiple 1 ml liquid cultures from different single colonies inside an anaerobic chamber. Once these cultures reached late exponential phase, they were used to inoculate separate 10 ml cultures with 10% v/v. Each reactor was then inoculated with 25% v/v from this liquid culture as it approached acceptor limitation (as it reached an OD_{600} of 0.48). Working electrodes were set at either -100 mV or +240 mV vs standard hydrogen electrode (SHE) and average current density recorded every 12 seconds. All experiments were carried out at 30 °C.

Stable isotope probe medium

Heavy isotope labeled medium was identical in chemical composition to normal growth medium, with an increase in the final heavy isotope of ^{15}N to 6 atom% in ammonium, ^{13}C to 6 atom% in both acetate carbons, and ^{2}H to 2 atom% in the water. Enriched isotopic chemicals were purchased from Cambridge

isotope ($^{15}NH_4Cl$ [NLM-467]) and Sigma (D_2O [151882], $^{13}CH_3{}^{13}CO_2Na$ [282014]). Increased strength labeled media for short term incubations was made using 3x the heavy isotope percentage with identical concentrations of chemical species.

Stable isotope probing

Once the desired biofilm growth stage had been reached, the potentiostat was paused, and medium was carefully exchanged for 15 mL of anaerobic labeled medium. For negative control samples, regular electrode growth medium was used instead of stable isotope enriched medium. Biofilms were then incubated for 6 hours with working electrodes poised at the same potential as during growth, or changed from -100 mV to +240 mV (Fig. 5F).

Fixation and embedding

After stable isotope incubation, graphite electrodes were harvested from each reactor, rinsed by placing in fresh medium once to remove planktonic cells, fixed at room temperature for 1 hr (2% glutaraldehyde, 50 mM HEPES, pH 6.8), and stored at 4 °C in fixing solution. Negative controls to test for non-specific isotope binding were rinsed and incubated for 6 hours in stable isotope enriched media before storing in fixative. Samples (still attached to graphite) were rinsed twice (50 mM HEPES, pH 7) before negative staining in 1% OsO_4, 50 mM HEPES, pH 7 for 2h. After rinsing again, samples were incubated for 1h in 1% uranyl-acetate and rinsed. Samples were then dehydrated with sequential 10 minute incubations in 25%, 50%, 75%, and 100% ETOH. A 50/50 v/v LR White resin (Sigma-Aldrich, uncatalyzed) and ETOH solution was used to incubate samples at room temperature overnight. Samples were then placed in 2 mL microcentrifuge tubes containing enough LR White solution (catalyzed with benzoyl peroxide) to cover ~0.5 cm above the graphite flag (~1.5 mL) and incubated at 60 °C overnight.

Sample preparation for nanoSIMS

Resin-embedded biofilm-covered electrodes were cut perpendicular to the largest face of the electrode with a jeweler's saw and a smooth face was prepared using a microtome and glass knife. Thin sections between 200-500nm were cut for nanoSIMS analysis with a diamond knife. Floated sections were collected on glow discharged 7mm x 7mm silicon wafers (Active Business Company GmbH). Sections on silicon wafers were coated with 40nm gold using a Cressington sputter coater.

Electron microscopy

A subset of biofilm sections were imaged with transmission and scanning electron microscopy. For TEM, 100nm sections were cut using a diamond knife and collected on copper TEM grids and imaged on an FEI Tecnai (T12) operated at 120 keV. Sections collected on silicon wafers (described above) were imaged on a Merlin Compact SEM (Zeiss).

NanoSIMS data acquisition

Isotope enrichment data was collected on a CAMECA nanoSIMS 50L housed in the Center for Microanalysis at the California Institute of Technology. Gold coated samples were pre-sputtered with a 90pA primary Cs^+ ion beam with detector D1=1 until the $^{14}N^{12}C^-$ ion counts stabilized. Data was collected using a 0.5 pA beam with D1=3, and ES=2. Six masses were collected corresponding to the $^1H^-$, $^2H^-$, $^{12}C^-$, $^{13}C^-$, $^{14}N^{12}C^-$, and $^{15}N^{12}C^-$ ions, for the determination of $^2H/^1H$, $^{13}C/^{12}C$ and $^{15}N/^{14}N$ ratios, respectively using a tuning similar to that described in Kopf *et al.*, (Kopf et al., 2015). Acquisitions were 10 to 40μm square raster areas, with 128x128 to 512x512 pixels, and 1-4 planes were collected per area.

Dwell time settings resulting in acquisition times from 25 minutes to two hours per plane depending on the size of the rasters.

Data processing

NanoSIMS .im data files were initially processed using the Look@nanoSIMS Matlab GUI (Polerecky et al., 2012), to align planes and export raw data. All subsequent data processing and analysis was done in Matlab. Regions of acquisitions that contained *Geobacter* biofilm were outlined on the $^{14}N^{12}C^-$ mass image, where the surface of the electrode was traced by hand, and each pixel of *Geobacter* biofilm was assigned a minimum distance to the electrode surface by calculating the pairwise distance between each pixel in the biofilm and the electrode surface. Biofilm pixels were assigned to bins of 0.5 μm increments from the anode surface, and the $^{15}N^{12}C^-$ and $^{14}N^{12}C^-$ counts were pooled for each distance bin to calculate the fractional abundance of the heavy isotopes: i.e. $^{15}F = {}^{15}N^{12}C^-/({}^{15}N^{12}C^- + {}^{14}N^{12}C^-)$. Pixels with low $^{14}N^{12}C$ counts corresponding to the epoxy resin in between cells were omitted, as was the furthest distance bin when it contained very few pixels.

Acknowledgements

This publication was supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research [DE – SC0016469] and the NASA Astrobiology Institute, award #NNA13AA92A (to VJO) and the Simons Foundation (Program: Life Sciences-Simons Collaboration on Principles of Microbial Ecosystems; Award Number: 542393). GLC was supported by the NIH/NRSA training grant, T32 GM007616. FJO was supported by the Mexican National Council for Science and Technology. We thank Dr. Gail Celio for training and loan of the UMN Imaging facilities,

Dr. Ryan Hunter for assistance developing a biofilm staining and embedding protocol, and Dr. Yunbin Guan for assistance with the nanoSIMS analysis.

Figures

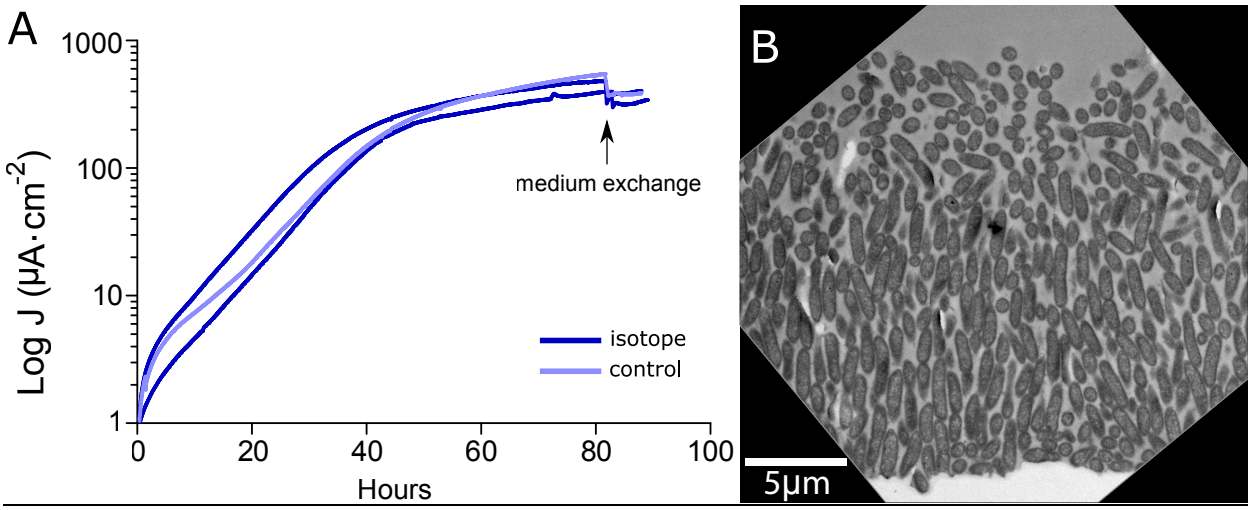


Figure 1 *G. sulfurreducens* biofilm structure is fully developed as the current density plateau is reached. Current density generation (**A**) and transmission electron microscopy images (**B**) of *G. sulfurreducens* biofilms. Cells were inoculated at time zero on anodes poised at +240 mV. After 83 hours of growth (see arrow), normal growth medium was exchanged with stable isotope labeled medium or unlabeled medium as a control.

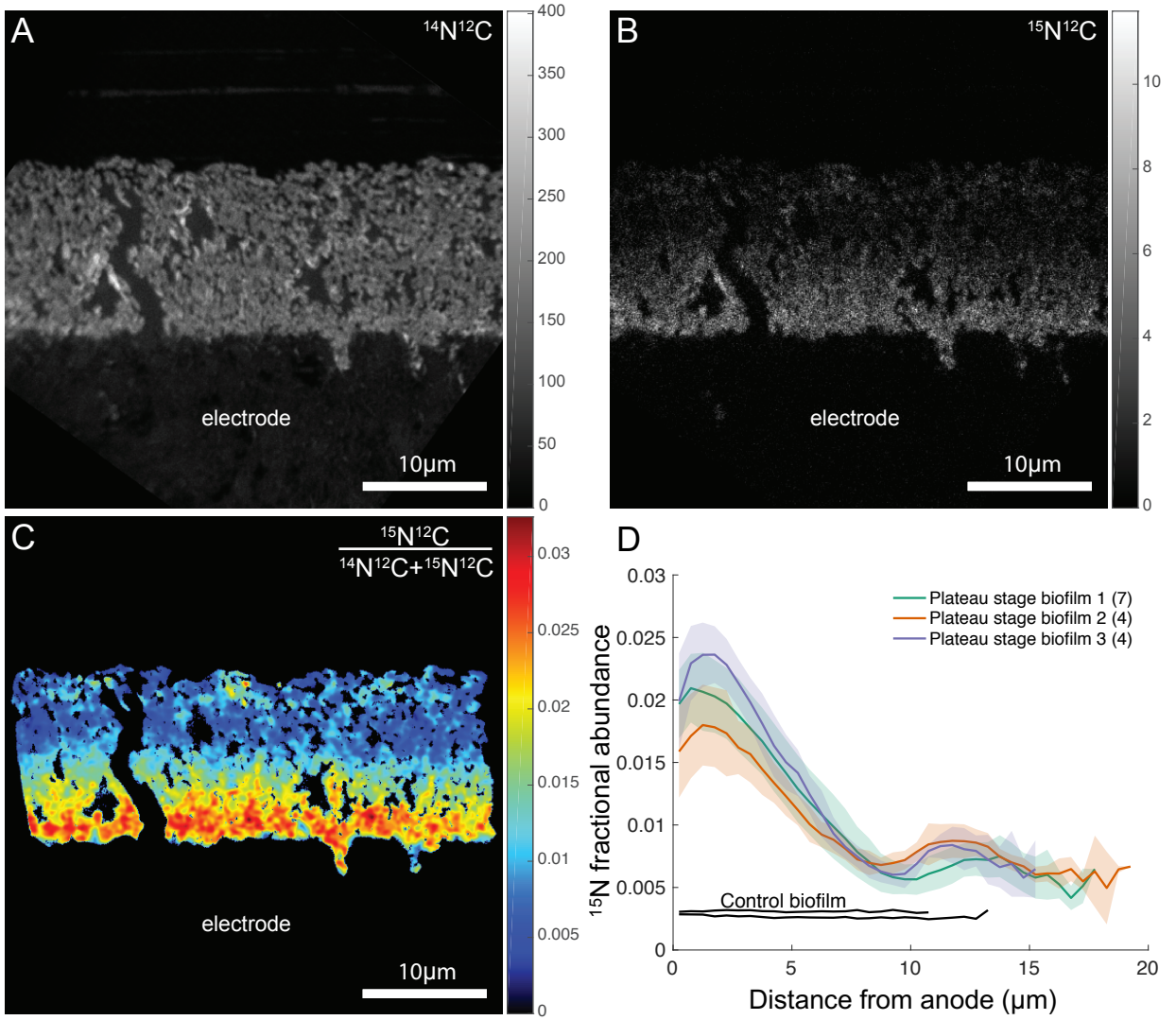


Figure 2 ^{15}N incorporation reveals anabolic activity patterns within *G. sulfurreducens* biofilms. Representative nanoSIMS analysis of *G. sulfurreducens* biofilms. (**A**) $^{14}N^{12}C^-$ ion image demonstrating biofilm morphology and spatial resolution. The graphite anode is located below the biofilm and the vertical black zone is a crack in the resin. Heavy nitrogen isotope image $^{15}N^{12}C^-$ (**B**) and ^{15}N fractional abundance image (**C**) reveal an enrichment at the anode surface. Graded scales to the right of **A**, **B**, and **C** serve as legend of abundance (**A** and **B**) or fractional abundance (**C**) measured throughout the biofilm. In (**C**) only pixels corresponding to the biofilm are displayed (see methods), and a 1 pixel radius gaussian blur is applied for image clarity. (**D**) Anabolic activity patterns in replicate *G. sulfurreducens* biofilms, n=3. Colored lines represent the average ^{15}N fractional abundance of different locations within a single biofilm. Transparent envelopes surrounding lines represent standard deviation of fractional abundance at each distance calculated based on multiple acquisitions of each biofilm (number of acquisitions per biological replicate are shown in parentheses in the legend). Black lines represent controls that were chemically fixed before incubating with isotopically labeled media, indicating live cells are required for isotope incorporation.

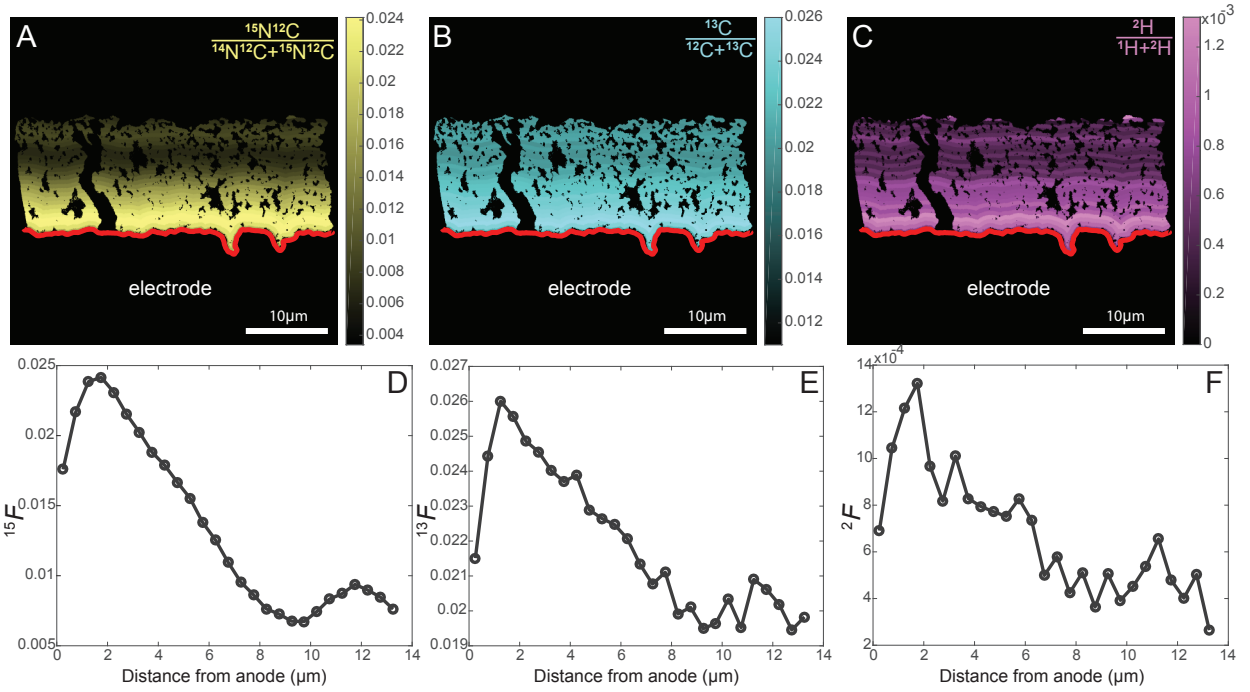


Figure 3 ^{13}C and ^{2}H incorporation follow the same pattern as ^{15}N. Fractional abundance of ^{15}N (**A**), ^{13}C (**B**) and ^{2}H (**C**) binned every half micron from the anode surface (red line). Graded scales to the right of **A**, **B**, and **C** serve as legend of fractional abundance measured throughout the biofilm. **D**-**F**) half-micron binned fractional abundance as a function of distance from electrode for ^{15}N, ^{13}C, and ^{2}H, respectively from representative samples. *F*= fractional abundance.

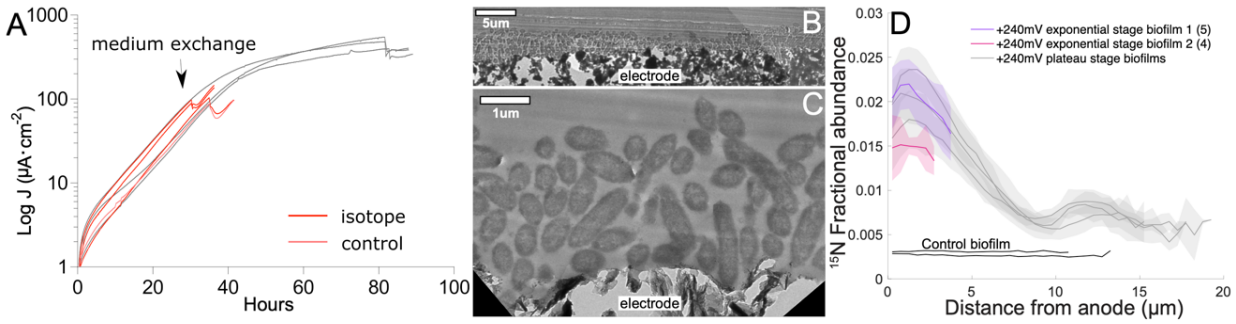


Figure 4 The decrease in anabolic activity with distance from the electrode is already present in thin biofilms. Current density (**A**) and transmission electron microscopy images (**B** and **C**) of representative *G. sulfurreducens* biofilms harvested during exponential phase. Graphite electrodes poised at +240 mV vs. SHE served as the terminal electron acceptor and normal growth media was exchanged with stable isotope labeled media after 35 hours of growth (arrow on **A**). (**D**) Anabolic activity patterns in duplicate *G. sulfurreducens* biofilms. Red and purple lines represent the average ^{15}N fractional abundance at different locations within a single biofilm (number of acquisitions per biological replicate are shown in parentheses in the legend). Transparent envelopes surrounding lines represent standard deviation of fractional abundance at each distance. Black lines represent controls that were chemically fixed before incubating with isotopically labeled media. For comparison, data from **A** and **D** is overlaid on data in gray from **Fig. 1A** and **Fig. 2D,** respectively, to show the similar maximum and decrease with distance features.

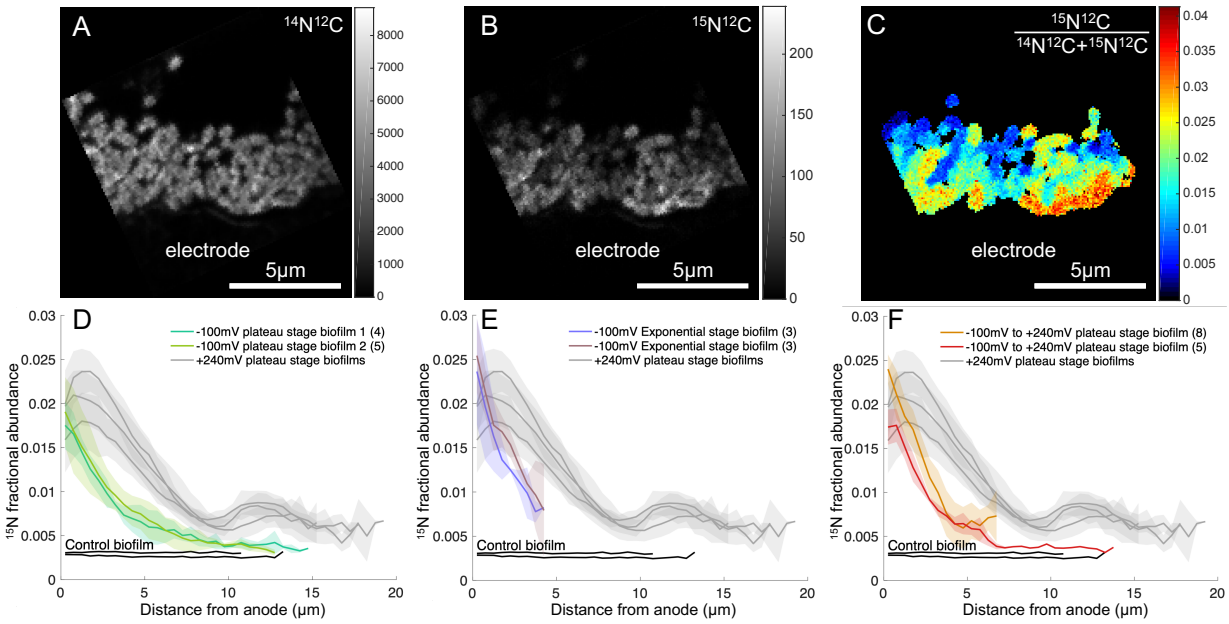


Figure 5 Lowering the poised potential of electrodes causes anabolic activity of *G. sulfurreducens* to decrease even more rapidly with distance, and cannot be restored by exposure to a more favorable reduction potential. Representative nanoSIMS analysis of *G. sulfurreducens* biofilms grown with graphite electrodes poised at -100 mV vs. SHE as sole terminal electron acceptor. (**A**) $^{14}N^{12}C^-$ ion image demonstrating biofilm morphology and spatial resolution. Graphite anode is located in the bottom, below the biofilm. Heavy nitrogen isotope image $^{15}N^{12}C^-$ (**B**) and ^{15}N fractional abundance image (**C**) reveal an enrichment at the anode surface. Graded scale to the right of **A**, **B**, and **C** serves as legend of fractional abundance measured throughout the biofilm. In (**C**) only pixels corresponding to the biofilm are displayed (see methods). Anabolic activity patterns in replicate *G. sulfurreducens* biofilms grown with graphite electrodes poised at -100 mV vs. SHE as sole terminal electron acceptor and harvested once current reached plateau (**D**), during exponential phase (**E**), or once current reached plateau and the reduction potential of the working electrode was raised to +240 mV vs. SHE for the duration of isotope label incorporation (**F**). Lines represent the average ^{15}N fractional abundance for analyses of different locations within a single biofilm (number of acquisitions per biological replicate are shown in parentheses in the legend). Transparent envelopes surrounding lines represent standard deviation of fractional abundance at each distance. Black lines represent controls that were chemically fixed before incubating with isotopically labeled medium. For comparison, data from **D**-**F** are overlaid on data in gray from **Fig. 2D**.

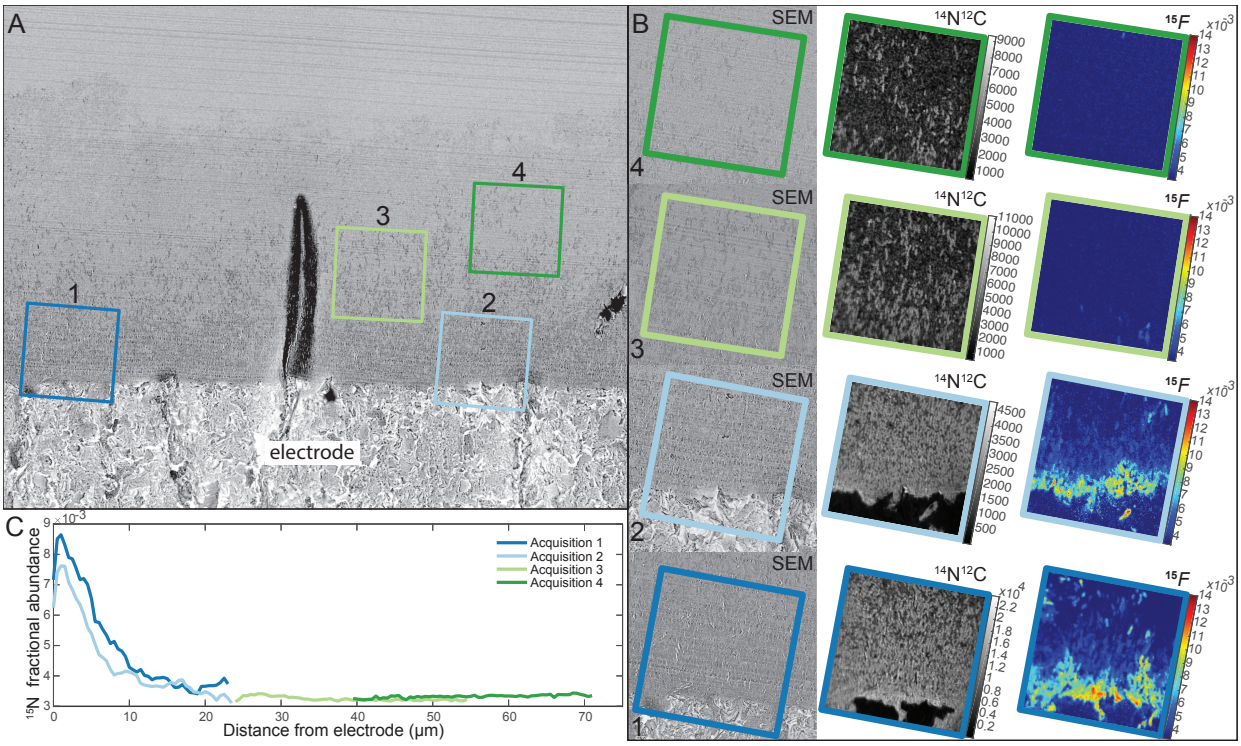


Figure 6 Anabolic activity pattern of *G. sulfurreducens* biofilms does not change after growth is maintained for two weeks. (**A**) Electron microscopy and nanoSIMS analysis of representative two-week old biofilm. Scanning electron microscopy image showing a two week old biofilm grown at +240 mV, demonstrating the thickness of the biofilm and regions of cell lysis in the upper layers. Representative nanoSIMS images were taken in regions 1-4. (**B**) The $^{14}N^{12}C^-$ ion image is shown along with the ^{15}N fractional abundance for each region next to their enlarged SEM image. (**C**) Anabolic activity patterns in two-week old *G. sulfurreducens* biofilms. Dark blue, light blue, light green, and dark green represent the ^{15}N fractional abundance for analyses of regions 1, 2, 3, and 4, respectively. Data representative of three biological replicates. Acquisition regions are 30 μm square and serve as scale bars.

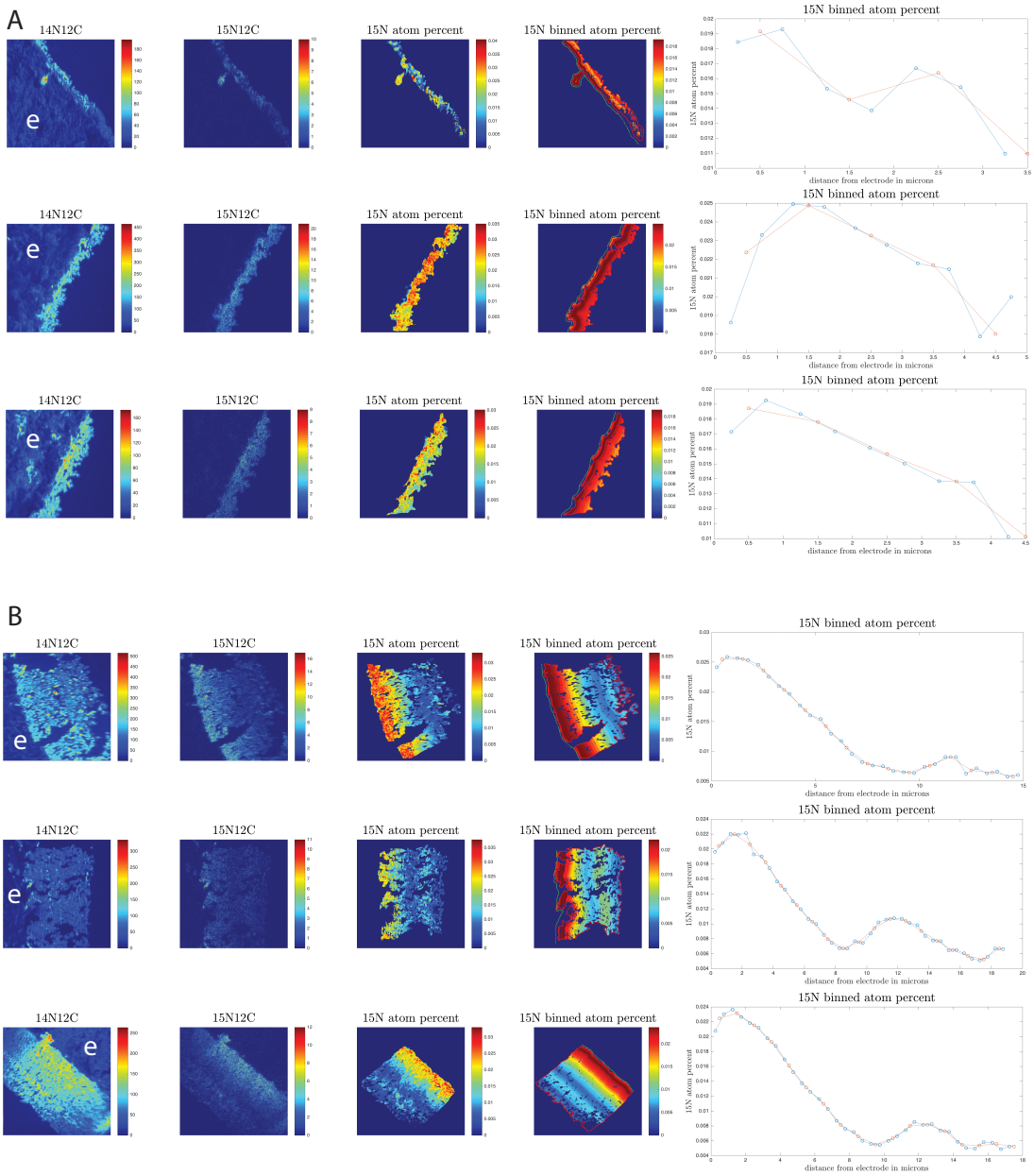


Figure S1 Repeatability of nanoSIMS analysis of *G. sulfurreducens* biofilms grown using a graphite electrode poised at +240 mV vs. SHE during (**A**) exponential and (**B**) current plateau phases. Panels from left to right correspond to $^{14}N^{12}C^-$ ion image demonstrating biofilm morphology and spatial resolution. The location of the graphite anode is labeled on each $^{14}N^{12}C$ ion panel with the letter "e". Second panels correspond to heavy nitrogen isotope image $^{15}N^{12}C^-$, and third panels to ^{15}N fractional abundance image, all revealing an enrichment at the anode surface. Plots to the right of nanoSIMS images correspond to binned anabolic activity patterns in each additional representative *G. sulfurreducens* biofilm. These data all contributed to the average activity plots in Figure 2D and Figure 4D.

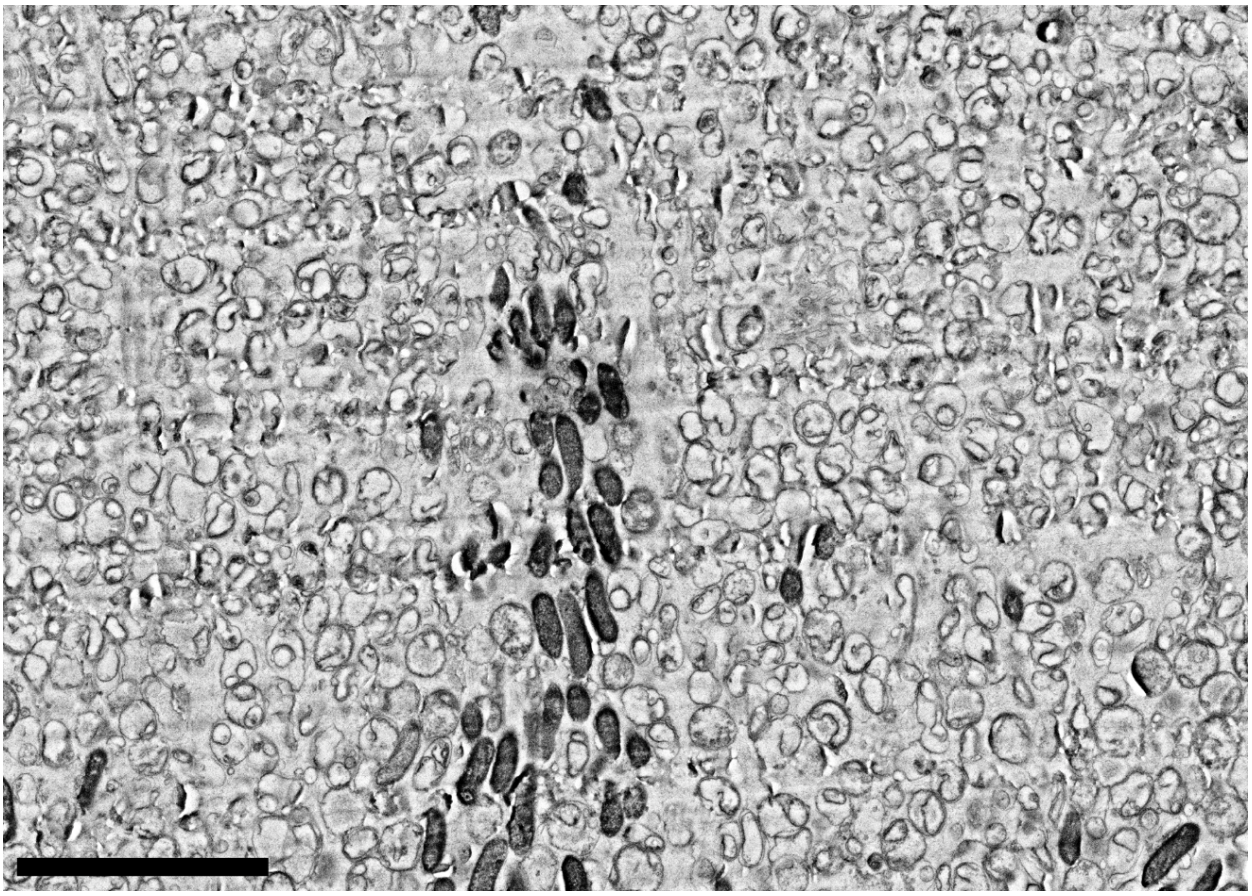


Figure S2 SEM image of a region of lysing cells in thin section of weeks old *G. sulfurreducens* biofilms (detail of biofilm from Fig. 6a in the main text). Cells with more electron density retained the characteristic size and shape of healthy *Geobacter* rods, while low density material appeared to be made up of lysed cells with irregular broken membranes. Intact cells yielded much greater $^{14}N^{12}C^-$ ion counts in nanoSIMS images, indicating dense biological material, and are oriented perpendicular to the surface of the anode (~20μm below the region displayed here). Scale bar is 5 μm.

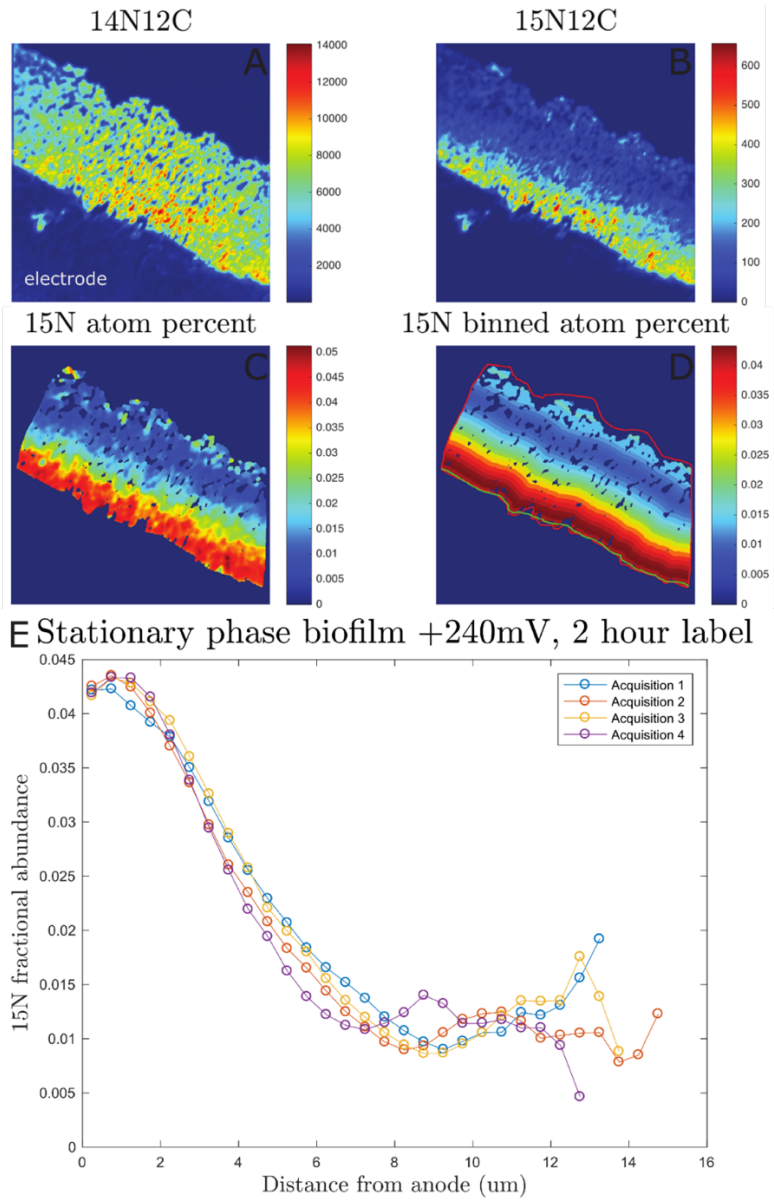


Figure S3 Short stable isotope probe experiment with higher percent heavy isotope label reveals similar activity patterns. (**A**-**D**) Representative nanoSIMS image of the biofilm grown at +240 mV and labeled for ~2 hours with 3x heavy isotope percent medium once current plateau was reached. Acquisition region 30x30µm. In (**C**) and (**D**) only regions defined as biofilm by manual tracing and with 14N12C counts above background are shown (as described in the methods). In (**D**) manual trace of the biofilm is shown in red, with electrode trace in green. (**E**) Binned fractional abundance of 15N in the biofilm at four different acquisition locations demonstrating the characteristic enriched activity at the anode interface and small peak in activity at the top of the biofilm.

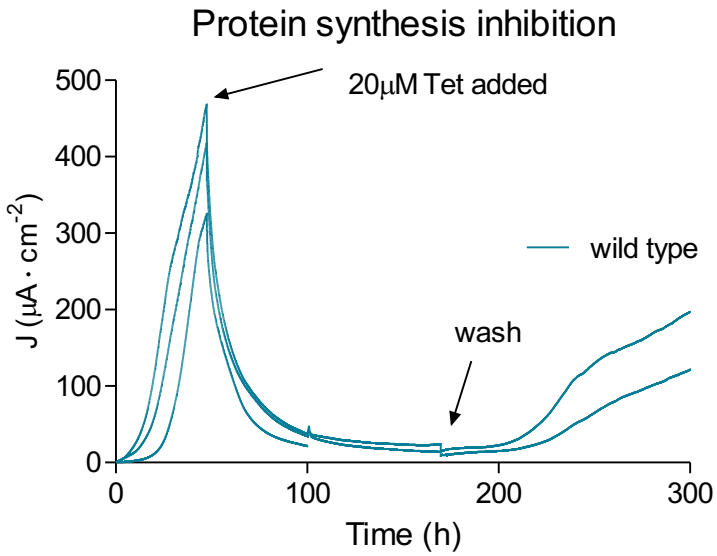


Figure S4 Current production of wild type *G. sulfurreducens* biofilms grown in triplicate until current reached plateau, then 20 μM tetracycline (Tet) was added. Once current density remained below 30 μA·cm-2 for 24 hours, medium was exchanged for fresh medium with no antibiotic.

References

Aleem MIH. 1966. Generation of reducing power in chemosynthesis II. Energy-linked reduction of pyridine nucleotides in the chemoautotroph, Nitrosomonas europaea. *Biochimica et Biophysica Acta (BBA) - Enzymology and Biological Oxidation* **113**:216–224. doi:10.1016/S0926-6593(66)80062-0

Babauta JT, Nguyen HD, Harrington TD, Renslow R, Beyenal H. 2012. pH, redox potential and local biofilm potential microenvironments within *Geobacter sulfurreducens* biofilms and their roles in electron transfer. *Biotechnol Bioeng* **109**:2651–2662. doi:10.1002/bit.24538

Baradaran R, Berrisford JM, Minhas GS, Sazanov LA. 2013. Crystal structure of the entire respiratory complex I. *Nature* **494**:443–448. doi:10.1038/nature11871

Battchikova N, Eisenhut M, Aro E-M. 2011. Cyanobacterial NDH-1 complexes: Novel insights and remaining puzzles. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, Regulation of Electron Transport in Chloroplasts **1807**:935–944. doi:10.1016/j.bbabio.2010.10.017

Battin TJ, Kaplan LA, Denis Newbold J, Hansen CME. 2003. Contributions of microbial biofilms to ecosystem processes in stream mesocosms. *Nature* **426**:439–442. doi:10.1038/nature02152

Bäumer S, Ide T, Jacobi C, Johann A, Gottschalk G, Deppenmeier U. 2000. The F420H2 Dehydrogenase fromMethanosarcina mazei Is a Redox-driven Proton Pump Closely Related to NADH Dehydrogenases. *J Biol Chem* **275**:17968–17973. doi:10.1074/jbc.M000650200

Belevich G, Knuuti J, Verkhovsky MI, Wikström M, Verkhovskaya M. 2011. Probing the mechanistic role of the long α-helix in subunit L of respiratory Complex I from Escherichia coli by site-directed mutagenesis. *Molecular Microbiology* **82**:1086–1095. doi:10.1111/j.1365-2958.2011.07883.x

Bonanni PS, Bradley DF, Schrott GD, Busalmen JP. 2013. Limitations for current production in *Geobacter sulfurreducens* biofilms. *ChemSusChem* **6**:711–720. doi:10.1002/cssc.201200671

Bond DR, Holmes DE, Tender LM, Lovley DR. 2002. Electrode-reducing microorganisms that harvest energy from marine sediments. *Science* **295**:483–485. doi:10.1126/science.1066771

Cerqueira NMFSA, Gonzalez PJ, Fernandes PA, Moura JJG, Ramos MJ. 2015. Periplasmic Nitrate Reductase and Formate Dehydrogenase: Similar Molecular Architectures with Very Different Enzymatic Activities. *Acc Chem Res* **48**:2875–2884. doi:10.1021/acs.accounts.5b00333

Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, Jehmlich N, Palatinszky M, Vierheilig J, Bulaev A, Kirkegaard RH, von Bergen M, Rattei T, Bendinger B, Nielsen PH, Wagner M. 2015. Complete nitrification by Nitrospira bacteria. *Nature* **528**:504–509. doi:10.1038/nature16461

de Beer D, Stoodley P, Roe F, Lewandowski Z. 1994. Effects of biofilm structures on oxygen distribution and mass transport. *Biotechnol Bioeng* **43**:1131–1138. doi:10.1002/bit.260431118

Deppenmeier U, Blaut M, Mahlmann A, Gottschalk G. 1990. Reduced coenzyme F420: heterodisulfide oxidoreductase, a proton- translocating redox system in methanogenic bacteria. *PNAS* **87**:9449–9453. doi:10.1073/pnas.87.23.9449

Deschamps P, Zivanovic Y, Moreira D, Rodriguez-Valera F, López-García P. 2014. Pangenome Evidence for Extensive Interdomain Horizontal Transfer Affecting Lineage Core and Shell Genes in Uncultured Planktonic Thaumarchaeota and Euryarchaeota. *Genome Biol Evol* **6**:1549–1563. doi:10.1093/gbe/evu127

Domozych DS, Domozych CR. 2008. Desmids and biofilms of freshwater wetlands: development and microarchitecture. *Microb Ecol* **55**:81–93. doi:10.1007/s00248-007-9253-y

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–1797. doi:10.1093/nar/gkh340

E. Franks A, P. Nevin K, Jia H, Izallalen M, L. Woodard T, R. Lovley D. 2009. Novel strategy for three-dimensional real-time imaging of microbial fuel cell communities: monitoring the inhibitory effects of proton accumulation within the anode biofilm. *Energy & Environmental Science* **2**:113–119. doi:10.1039/B816445B

Efremov RG, Baradaran R, Sazanov LA. 2010. The architecture of respiratory complex I. *Nature* **465**:441–445. doi:10.1038/nature09066

Efremov RG, Sazanov LA. 2012. The coupling mechanism of respiratory complex I — A structural and evolutionary perspective. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 17th European Bioenergetics Conference **1817**:1785–1795. doi:10.1016/j.bbabio.2012.02.015

Ehrich S, Behrens D, Lebedeva E, Ludwig W, Bock E. 1995. A new obligately chemolithoautotrophic, nitrite-oxidizing bacterium,Nitrospira moscoviensis sp. nov. and its phylogenetic relationship. *Arch Microbiol* **164**:16–23. doi:10.1007/BF02568729

Ehrlich HL. 2016. Uppermost lithosphere as a microbial habitatEhrlich's Geomicrobiology. Boca Raton, FL: CRC Press. pp. 55–68.

Fischer WW, Hemp J, Johnson JE. 2016. Evolution of Oxygenic Photosynthesis. *Annual Review of Earth and Planetary Sciences* **44**:647–683. doi:10.1146/annurev-earth-060313-054810

Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. 2016. Biofilms: an emergent form of bacterial life. *Nat Rev Micro* **14**:563–575. doi:10.1038/nrmicro.2016.94

Franks AE, Glaven RH, Lovley DR. 2012a. Real-time spatial gene expression analysis within current-producing biofilms. *ChemSusChem* **5**:1092–1098. doi:10.1002/cssc.201100714

Franks AE, Glaven RH, Lovley DR. 2012b. Real-Time Spatial Gene Expression Analysis within Current-Producing Biofilms. *ChemSusChem* **5**:1092–1098. doi:10.1002/cssc.201100714

Franks AE, Nevin KP, Glaven RH, Lovley DR. 2010. Microtoming coupled to microarray analysis to evaluate the spatial metabolic status of *Geobacter sulfurreducens* biofilms. *The ISME Journal* **4**:509. doi:10.1038/ismej.2009.137

Friedrich T. 1998. The NADH:ubiquinone oxidoreductase (complex I) from Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1364**:134–146. doi:10.1016/S0005-2728(98)00024-3

Friedrich T, Scheide D. 2000. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Letters* **479**:1–5. doi:10.1016/S0014-5793(00)01867-6

Friedrich T, Steinmüller K, Weiss H. 1995. The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts. *FEBS Letters* **367**:107–111. doi:10.1016/0014-5793(95)00548-N

Gantzer CJ, Rittmann BE, Herricks EE. 1988. Mass transport to streambed biofilms. *Water Research* **22**:709–722. doi:10.1016/0043-1354(88)90182-0

Gao F, Zhao J, Wang X, Qin S, Wei L, Ma W. 2016. NdhV Is a Subunit of NADPH Dehydrogenase Essential for Cyclic Electron Transport in Synechocystis sp. Strain PCC 6803. *Plant Physiology* **170**:752–760. doi:10.1104/pp.15.01430

Grimaldi S, Schoepp-Cothenet B, Ceccaldi P, Guigliarelli B, Magalon A. 2013. The prokaryotic Mo/W-bisPGD enzymes family: A catalytic workhorse in bioenergetic. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, Metals in Bioenergetics and Biomimetics Systems **1827**:1048–1085. doi:10.1016/j.bbabio.2013.01.011

Hallberg ZF, Chan CH, Wright TA, Kranzusch PJ, Doxzen KW, Park JJ, Bond DR, Hammond MC. n.d. Structure and mechanism of a Hypr GGDEF enzyme that activates cGAMP signaling to control extracellular metal respiration. *eLife* **8**. doi:10.7554/eLife.43959

Hicks DB, Liu J, Fujisawa M, Krulwich TA. 2010. F1F0-ATP synthases of alkaliphilic bacteria: Lessons from their adaptations. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1797**:1362–1377. doi:10.1016/j.bbabio.2010.02.028

Holmes DE, Bond DR, O'Neil RA, Reimers CE, Tender LR, Lovley DR. 2004. Microbial communities associated with electrodes harvesting electricity from a variety of aquatic sediments. *Microb Ecol* **48**:178–190. doi:10.1007/s00248-003-0004-4

Hunte C, Zickermann V, Brandt U. 2010. Functional Modules and Structural Basis of Conformational Coupling in Mitochondrial Complex I. *Science* **329**:448–451. doi:10.1126/science.1191046

Inoue K, Leang C, Franks AE, Woodard TL, Nevin KP, Lovley DR. 2011. Specific localization of the *c*-type cytochrome OmcZ at the anode surface in current-producing biofilms of *Geobacter sulfurreducens*. *Environ Microbiol Rep* **3**:211–217. doi:10.1111/j.1758-2229.2010.00210.x

Islam FS, Gault AG, Boothman C, Polya DA, Charnock JM, Chatterjee D, Lloyd JR. 2004. Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* **430**:68–71. doi:10.1038/nature02638

Jain A, Gazzola G, Panzera A, Zanoni M, Marsili E. 2011. Visible spectroelectrochemical characterization of *Geobacter sulfurreducens* biofilms on optically transparent indium tin oxide electrode. *Electrochimica Acta*, Selected Papers from the 61st ISE Meeting, Nice, France, 2010 **56**:10776–10785. doi:10.1016/j.electacta.2011.02.073

Jiménez Otero F, Chan CH, Bond DR. 2018. Identification of different putative outer membrane electron conduits necessary for Fe(III) citrate, Fe(III) oxide, Mn(IV) oxide, or electrode reduction by *Geobacter sulfurreducens*. *Journal of Bacteriology* **200**:e00347-18. doi:10.1128/JB.00347-18

Jones AJY, Blaza JN, Varghese F, Hirst J. 2017. Respiratory Complex I in Bos taurus and Paracoccus denitrificans Pumps Four Protons across the Membrane for Every NADH Oxidized. *J Biol Chem* **292**:4987–4995. doi:10.1074/jbc.M116.771899

Jung S, Regan JM. 2007. Comparison of anode bacterial communities and performance in microbial fuel cells with different electron donors. *Appl Microbiol Biotechnol* **77**:393–402. doi:10.1007/s00253-007-1162-y

Kato Marcus A, Torres CI, Rittmann BE. 2007. Conduction-based modeling of the biofilm anode of a microbial fuel cell. *Biotechnol Bioeng* **98**:1171–1182. doi:10.1002/bit.21533

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**:3059–3066. doi:10.1093/nar/gkf436

Kirchhoff C, Cypionka H. 2017. Propidium ion enters viable cells with high membrane potential during live-dead staining. *Journal of Microbiological Methods* **142**:79–82. doi:10.1016/j.mimet.2017.09.011

Koch H, Lücker S, Albertsen M, Kitzinger K, Herbold C, Spieck E, Nielsen PH, Wagner M, Daims H. 2015. Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *PNAS* **112**:11371–11376. doi:10.1073/pnas.1506533112

Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**:543–546. doi:10.1038/nature03911

Könneke M, Schubert DM, Brown PC, Hügler M, Standfest S, Schwander T, Borzyskowski LS von, Erb TJ, Stahl DA, Berg IA. 2014. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *PNAS* **111**:8239–8244. doi:10.1073/pnas.1402028111

Kopf SH, McGlynn SE, Green-Saxena A, Guan Y, Newman DK, Orphan VJ. 2015. Heavy water and 15N labelling with NanoSIMS analysis reveals growth rate-dependent metabolic heterogeneity in chemostats. *Environmental Microbiology* **17**:2542–2556. doi:10.1111/1462-2920.12752

Kopf SH, Sessions AL, Cowley ES, Reyes C, Van Sambeek L, Hu Y, Orphan VJ, Kato R, Newman DK. 2016. Trace incorporation of heavy water reveals slow and heterogeneous pathogen growth rates in cystic fibrosis sputum. *Proc Natl Acad Sci USA* **113**:E110-116. doi:10.1073/pnas.1512057112

Lebedev N, Strycharz-Glaven SM, Tender LM. 2014. Spatially resolved confocal resonant Raman microscopic analysis of anode-grown *Geobacter sulfurreducens* biofilms. *ChemPhysChem* **15**:320–327. doi:10.1002/cphc.201300984

Lee H-S, Parameswaran P, Kato-Marcus A, Torres CI, Rittmann BE. 2008. Evaluation of energy-conversion efficiencies in microbial fuel cells (MFCs) utilizing fermentable and non-fermentable substrates. *Water Res* **42**:1501–1510. doi:10.1016/j.watres.2007.10.036

Lemaire R, Webb RI, Yuan Z. 2008. Micro-scale observations of the structure of aerobic microbial granules used for the treatment of nutrient-rich industrial wastewater. *ISME J* **2**:528–541. doi:10.1038/ismej.2008.12

Levar CE, Chan CH, Mehta-Kolte MG, Bond DR. 2014. An inner membrane cytochrome required only for reduction of high redox potential extracellular electron acceptors. *mBio* **5**:e02034-14. doi:10.1128/mBio.02034-14

Levar CE, Hoffman CL, Dunshee AJ, Toner BM, Bond DR. 2017. Redox potential as a master variable controlling pathways of metal reduction by *Geobacter sulfurreducens*. *The ISME Journal* **11**:741. doi:10.1038/ismej.2016.146

Liu Y, Kim H, Franklin RR, Bond DR. 2011. Linking spectral and electrochemical analysis to monitor *c*-type cytochrome redox status in living *Geobacter sulfurreducens* biofilms. *ChemPhysChem* **12**:2235–2241. doi:10.1002/cphc.201100246

Lücker S, Nowka B, Rattei T, Spieck E, Daims H. 2013. The Genome of Nitrospina gracilis Illuminates the Metabolism and Evolution of the Major Marine Nitrite Oxidizer. *Front Microbiol* **4**. doi:10.3389/fmicb.2013.00027

Lücker S, Wagner M, Maixner F, Pelletier E, Koch H, Vacherie B, Rattei T, Damsté JSS, Spieck E, Paslier DL, Daims H. 2010. A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *PNAS* **107**:13479–13484. doi:10.1073/pnas.1003860107

Markowitz VM, Chen I-MA, Chu K, Szeto E, Palaniappan K, Jacob B, Ratner A, Liolios K, Pagani I, Huntemann M, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG/M-HMP: A Metagenome Comparative Analysis System for the Human Microbiome Project. *PLOS ONE* **7**:e40151. doi:10.1371/journal.pone.0040151

Marsili E, Rollefson JB, Baron DB, Hozalski RM, Bond DR. 2008a. Microbial biofilm voltammetry: direct electrochemical characterization of catalytic electrode-attached biofilms. *Appl Environ Microbiol* **74**:7329–7337. doi:10.1128/AEM.00177-08

Marsili E, Rollefson JB, Baron DB, Hozalski RM, Bond DR. 2008b. Microbial Biofilm Voltammetry: Direct Electrochemical Characterization of Catalytic Electrode-Attached Biofilms. *Appl Environ Microbiol* **74**:7329–7337. doi:10.1128/AEM.00177-08

Marsili E, Sun J, Bond DR. 2010. Voltammetry and growth physiology of *Geobacter sulfurreducens* biofilms as a function of growth stage and imposed electrode potential. *Electroanalysis* **22**:865–874. doi:10.1002/elan.200800007

McGlynn SE, Chadwick GL, Kempes CP, Orphan VJ. 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* **526**:531–535. doi:10.1038/nature15512

Meier T, Morgner N, Matthies D, Pogoryelov D, Keis S, Cook GM, Dimroth P, Brutschy B. 2007. A tridecameric c ring of the adenosine triphosphate (ATP) synthase from the thermoalkaliphilic Bacillus sp. strain TA2.A1 facilitates ATP synthesis at low electrochemical proton potential. *Molecular Microbiology* **65**:1181–1192. doi:10.1111/j.1365-2958.2007.05857.x

Mitchell P. 1966. Chemiosmotic Coupling in Oxidative and Photosynthetic Phosphorylation. *Biological Reviews* **41**:445–501. doi:10.1111/j.1469-185X.1966.tb01501.x

Moparthi VK, Hägerhäll C. 2011. The Evolution of Respiratory Chain Complex I from a Smaller Last Common Ancestor Consisting of 11 Protein Subunits. *J Mol Evol* **72**:484–497. doi:10.1007/s00239-011-9447-2

Moparthi VK, Kumar B, Al-Eryani Y, Sperling E, Górecki K, Drakenberg T, Hägerhäll C. 2014. Functional role of the MrpA- and MrpD-homologous protein subunits in enzyme complexes evolutionary related to respiratory chain complex I. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**:178–185. doi:10.1016/j.bbabio.2013.09.012

Nevin KP, Richter H, Covalla SF, Johnson JP, Woodard TL, Orloff AL, Jia H, Zhang M, Lovley DR. 2008. Power output and columbic efficiencies from biofilms of *Geobacter sulfurreducens* comparable to mixed community microbial fuel cells. *Environ Microbiol* **10**:2505–2514. doi:10.1111/j.1462-2920.2008.01675.x

Niftrik V, A L, Fuerst JA, Damsté JSS, Kuenen JG, Jetten MSM, Strous M. 2004. The anammoxosome: an intracytoplasmic compartment in anammox bacteria. *FEMS Microbiol Lett* **233**:7–13. doi:10.1016/j.femsle.2004.01.044

Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, Mclean JS, Lasken R, Tesler G,

Alekseyev MA, Pevzner PA. 2013. Assembling Single-Cell Genomes and Mini-Metagenomes From Chimeric MDA Products. *Journal of Computational Biology* **20**:714–737. doi:10.1089/cmb.2013.0084

Polerecky L, Adam B, Milucka J, Musat N, Vagner T, Kuypers MMM. 2012. Look@NanoSIMS--a tool for the analysis of nanoSIMS data in environmental microbiology. *Environ Microbiol* **14**:1009–1023. doi:10.1111/j.1462-2920.2011.02681.x

Preiss L, Klyszejko AL, Hicks DB, Liu J, Fackelmayer OJ, Yildiz Ö, Krulwich TA, Meier T. 2013. The c-ring stoichiometry of ATP synthase is adapted to cell physiological requirements of alkaliphilic Bacillus pseudofirmus OF4. *PNAS* **110**:7874–7879. doi:10.1073/pnas.1303333110

Preiss L, Langer JD, Hicks DB, Liu J, Yildiz Ö, Krulwich TA, Meier T. 2014. The c-ring ion binding site of the ATP synthase from Bacillus pseudofirmus OF4 is adapted to alkaliphilic lifestyle. *Molecular Microbiology* **92**:973–984. doi:10.1111/mmi.12605

Ramsing NB, Kühl M, Jørgensen BB. 1993. Distribution of sulfate-reducing bacteria, O_2, and H_2S in photosynthetic biofilms determined by oligonucleotide probes and microelectrodes. *Appl Environ Microbiol* **59**:3840–3849.

Reguera G, Nevin KP, Nicoll JS, Covalla SF, Woodard TL, Lovley DR. 2006. Biofilm and nanowire production leads to increased current in *Geobacter sulfurreducens* fuel cells. *Appl Environ Microbiol* **72**:7345–7348. doi:10.1128/AEM.01444-06

Renslow RS, Babauta JT, Dohnalkova AC, Boyanov MI, Kemner KM, Majors PD, Fredrickson JK, Beyenal H. 2013. Metabolic spatial variability in electrode-respiring Geobacter sulfurreducens biofilms. *Energy Environ Sci* **6**:1827–1836. doi:10.1039/C3EE40203G

Robuschi L, Tomba JP, Busalmen JP. 2017. Proving *Geobacter* biofilm connectivity with confocal Raman microscopy. *Journal of Electroanalytical Chemistry*, SI:Professor Antonio Aldaz **793**:99–103. doi:10.1016/j.jelechem.2016.11.005

Robuschi L, Tomba JP, Schrott GD, Bonanni PS, Desimone PM, Busalmen JP. 2013. Spectroscopic slicing to reveal internal redox gradients in electricity-producing biofilms. *Angew Chem Int Ed* **52**:925–928. doi:10.1002/anie.201205440

Rollefson JB, Stephen CS, Tien M, Bond DR. 2011. Identification of an extracellular polysaccharide network essential for cytochrome anchoring and biofilm formation in *Geobacter sulfurreducens*. *J Bacteriol* **193**:1023–1033. doi:10.1128/JB.01092-10

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protocols* **5**:725–738. doi:10.1038/nprot.2010.5

Rupprecht A, Sokolenko EA, Beck V, Ninnemann O, Jaburek M, Trimbuch T, Klishin SS, Jezek P, Skulachev VP, Pohl EE. 2010. Role of the Transmembrane Potential in the Membrane Proton Leak. *Biophysical Journal* **98**:1503–1511. doi:10.1016/j.bpj.2009.12.4301

Russell JB, Cook GM. 1995. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol Mol Biol Rev* **59**:48–62.

Sauer K, Camper AK, Ehrlich GD, Costerton JW, Davies DG. 2002. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J Bacteriol* **184**:1140–1154.

Sazanov LA. 2015. A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nature Reviews Molecular Cell Biology* **16**:nrm3997. doi:10.1038/nrm3997

Schmitz O, Bothe H. 1996. The diaphorase subunit HoxU of the bidirectional hydrogenase as electron transferring protein in cyanobacterial respiration? *Naturwissenschaften* **83**:525–527. doi:10.1007/BF01141957

Schrott GD, Ordoñez MV, Robuschi L, Busalmen JP. 2014. Physiological stratification in electricity-producing biofilms of *Geobacter sulfurreducens*. *ChemSusChem* **7**:598–603. doi:10.1002/cssc.201300605

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**:539. doi:10.1038/msb.2011.75

Simon J, Klotz MG. 2013. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, The evolutionary aspects of bioenergetic systems **1827**:114–135. doi:10.1016/j.bbabio.2012.07.005

Sleutels THJA, Ter Heijne A, Buisman CJN, Hamelers HVM. 2012. Bioelectrochemical systems: an outlook for practical applications. *ChemSusChem* **5**:1012–1019. doi:10.1002/cssc.201100732

Snider RM, Strycharz-Glaven SM, Tsoi SD, Erickson JS, Tender LM. 2012. Long-range electron transport in *Geobacter sulfurreducens* biofilms is redox gradient-driven. *PNAS* **109**:15467–15472. doi:10.1073/pnas.1209829109

Sorokin DY, Tourova TP, Kolganova TV, Detkova EN, Galinski EA, Muyzer G. 2011. Culturable diversity of lithotrophic haloalkaliphilic sulfate-reducing bacteria in soda lakes and the description of Desulfonatronumthioautotrophicum sp. nov., Desulfonatronum thiosulfatophilum sp. nov., Desulfonatronovibrio thiodismutans sp. nov., and Desulfonatronovibrio magnus sp. nov. *Extremophiles* **15**:391–401. doi:10.1007/s00792-011-0370-7

Sorokin DYu, Tourova TP, Henstra AM, Stams AJM, Galinski EA, Muyzer G. 2008. Sulfidogenesis under extremely haloalkaline conditions by Desulfonatronospira thiodismutans gen. nov., sp. nov., and Desulfonatronospira delicata sp. nov. – a novel lineage of Deltaproteobacteria from hypersaline soda lakes. *Microbiology* **154**:1444–1453. doi:10.1099/mic.0.2007/015628-0

Spang A, Poehlein A, Offre P, Zumbrägel S, Haider S, Rychlik N, Nowka B, Schmeisser C, Lebedeva EV, Rattei T, Böhm C, Schmid M, Galushko A, Hatzenpichler R, Weinmaier T, Daniel R, Schleper C, Spieck E, Streit W, Wagner M. 2012. The genome of the ammonia-oxidizing CandidatusNitrososphaera gargensis: insights into metabolic versatility and environmental adaptations. *Environ Microbiol* **14**:3122–3145. doi:10.1111/j.1462-2920.2012.02893.x

Steidl RJ, Lampa-Pastirk S, Reguera G. 2016. Mechanistic stratification in electroactive biofilms of *Geobacter sulfurreducens* mediated by pilus nanowires. *Nat Commun* **7**:12217. doi:10.1038/ncomms12217

Stephen CS, LaBelle EV, Brantley SL, Bond DR. 2014a. Abundance of the multiheme *c*-type cytochrome OmcB increases in outer biofilm layers of electrode-grown *Geobacter sulfurreducens*. *PLOS ONE* **9**:e104336. doi:10.1371/journal.pone.0104336

Stephen CS, LaBelle EV, Brantley SL, Bond DR. 2014b. Abundance of the Multiheme c-Type Cytochrome OmcB Increases in Outer Biofilm Layers of Electrode-Grown Geobacter sulfurreducens. *PLoS ONE* **9**:e104336. doi:10.1371/journal.pone.0104336

Stevenson BS, Eichorst SA, Wertz JT, Schmidt TM, Breznak JA. 2004. New Strategies for Cultivation and Detection of Previously Uncultured Microbes. *Appl Environ Microbiol* **70**:4748–4755. doi:10.1128/AEM.70.8.4748-4755.2004

Stieglmeier M, Klingl A, Alves RJE, Rittmann SK-MR, Melcher M, Leisch N, Schleper C. 2014. Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *International Journal of Systematic and Evolutionary Microbiology* **64**:2738–2752. doi:10.1099/ijs.0.063172-0

Stoodley P, Dodds I, Boyle JD, Lappin-Scott HM. 1998. Influence of hydrodynamics and nutrients on biofilm structure. *J Appl Microbiol* **85 Suppl 1**:19S-28S. doi:10.1111/j.1365-2672.1998.tb05279.x

Stoodley P, Sauer K, Davies DG, Costerton JW. 2002. Biofilms as complex differentiated communities. *Annu Rev Microbiol* **56**:187–209. doi:10.1146/annurev.micro.56.012302.160705

Summers ZM, Fogarty HE, Leang C, Franks AE, Malvankar NS, Lovley DR. 2010. Direct exchange of electrons within aggregates of an evolved syntrophic coculture of anaerobic bacteria. *Science* **330**:1413–1415. doi:10.1126/science.1196526

Sun D, Chen J, Huang H, Liu W, Ye Y, Cheng S. 2016. The effect of biofilm thickness on electrochemical activity of *Geobacter sulfurreducens*. *International Journal of Hydrogen Energy*, Special Issue: Hydrogen and Fuel Cell Developments: A special issue on the 8th International Conference on Sustainable Energy and Environmental Protection (SEEP 2015), 11–14 August 2015, Paisley, Scotland, UK **41**:16523–16528. doi:10.1016/j.ijhydene.2016.04.163

Sun D, Cheng S, Wang A, Li F, Logan BE, Cen K. 2015. Temporal-spatial changes in viabilities and electrochemical properties of anode biofilms. *Environ Sci Technol* **49**:5227–5235. doi:10.1021/acs.est.5b00175

Sun D, Cheng S, Zhang F, Logan BE. 2017. Current density reversibly alters metabolic spatial structure of exoelectrogenic anode biofilms. *Journal of Power Sources* **356**:566–571. doi:10.1016/j.jpowsour.2016.11.115

Tadanier CJ, Schreiber ME, Roller JW. 2005. Arsenic mobilization through microbially mediated deflocculation of ferrihydrite. *Environ Sci Technol* **39**:3061–3068. doi:10.1021/es048206d

Talavera G, Castresana J, Kjer K, Page R, Sullivan J. 2007. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst Biol* **56**:564–577. doi:10.1080/10635150701472164

Teal TK, Lies DP, Wold BJ, Newman DK. 2006. Spatiometabolic stratification of *Shewanella oneidensis* biofilms. *Appl Environ Microbiol* **72**:7324–7330. doi:10.1128/AEM.01163-06

Tejedor Sanz S, Fernández Labrador P, Hart S, Torres CI, Esteve-Núñez A. 2018. *Geobacter* dominates the inner layers of a stratified biofilm on a fluidized anode during brewery wastewater treatment. *Front Microbiol* **9**:387. doi:10.3389/fmicb.2018.00378

Tender LM, Reimers CE, Stecher HA, Holmes DE, Bond DR, Lowy DA, Pilobello K, Fertig SJ, Lovley DR. 2002. Harnessing microbially generated power on the seafloor. *Nat Biotechnol* **20**:821–825. doi:10.1038/nbt716

Torres CI, Kato Marcus A, Rittmann BE. 2008. Proton transport inside the biofilm limits electrical current generation by anode-respiring bacteria. *Biotechnol Bioeng* **100**:872–881. doi:10.1002/bit.21821

Torres-Bacete J, Sinha PK, Matsuno-Yagi A, Yagi T. 2011. Structural Contribution of C-terminal Segments of NuoL (ND5) and NuoM (ND4) Subunits of Complex I from Escherichia coli. *J Biol Chem* **286**:34007–34014. doi:10.1074/jbc.M111.260968

Valentine DL. 2007. Opinion: Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology* **5**:nrmicro1619. doi:10.1038/nrmicro1619

Vinothkumar KR, Zhu J, Hirst J. 2014. Architecture of mammalian respiratory complex I. *Nature* **515**:80–84. doi:10.1038/nature13686

Waite DW, Vanwonterghem I, Rinke C, Parks DH, Zhang Y, Takai K, Sievert SM, Simon J, Campbell BJ, Hanson TE, Woyke T, Klotz MG, Hugenholtz P. 2017. Comparative Genomic Analysis of the Class Epsilonproteobacteria and Proposed Reclassification to Epsilonbacteraeota (phyl. nov.). *Front Microbiol* **8**. doi:10.3389/fmicb.2017.00682

Wang F, Gu Y, O'Brian JP, Yi SM, Yalcin SE, Srikanth V, Shen C, Vu D, Ing NL, Hochbaum AI, Egelman EE, Malvankar NS. 2019. Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers. *Cell* **177**:361–369.

Weerakoon DR, Olson JW. 2008. The Campylobacter jejuni NADH:Ubiquinone Oxidoreductase (Complex I) Utilizes Flavodoxin Rather than NADH. *J Bacteriol* **190**:915–925. doi:10.1128/JB.01647-07

Wegener G, Krukenberg V, Riedel D, Tegetmeyer HE, Boetius A. 2015. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**:587–590. doi:10.1038/nature15733

Welte C, Deppenmeier U. 2014. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 18th European Bioenergetics Conference 2014 Lisbon, Portugal **1837**:1130–1147. doi:10.1016/j.bbabio.2013.12.002

Welte C, Deppenmeier U. 2011. Membrane-Bound Electron Transport in Methanosaeta thermophila. *J Bacteriol* **193**:2868–2870. doi:10.1128/JB.00162-11

Wilmes P, Remis JP, Hwang M, Auer M, Thelen MP, Banfield JF. 2009. Natural acidophilic biofilm communities reflect distinct organismal and functional organization. *The ISME Journal* **3**:266–270. doi:10.1038/ismej.2008.90

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Meth* **12**:7–8. doi:10.1038/nmeth.3213

Zacharoff L, Chan CH, Bond DR. 2016. Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of *Geobacter sulfurreducens*. *Bioelectrochemistry* **107**:7–13. doi:10.1016/j.bioelechem.2015.08.003

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**:40. doi:10.1186/1471-2105-9-40

Chapter 3

3. COMPARATIVE GENOMICS REVEALS ELECTRON TRANSFER AND SYNTROPHIC MECHANISMS DIFFERENTIATING METHANOTROPHIC AND METHANOGENIC ARCHAEA

Grayson L Chadwick[1*], Connor T Skennerton[1*], Rafael Laso-Pérez[2,3], Andy O Leu[4], Daan R Speth[1], Hang Yu[1], Connor Morgan-Lang[5], Roland Hatzenpichler[1,$], Danielle Goudeau[6], Rex Malmstrom[6], William J Brazelton[7], Tanja Woyke[6], Steven J Hallam[5,8,9,10], Gene W Tyson[4], Gunter Wegener[2,3], Antje Boetius[2,3,11], Victoria J Orphan[1]

1. Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA 91125
2. Max-Planck Institute for Marine Microbiology, 28359 Bremen, Germany
3. MARUM, Center for Marine Environmental Science, 28359 and Department of Geosciences, University of Bremen, Bremen, Germany
4. Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of Queensland, Brisbane, Queensland, Australia
5. Graduate Program in Bioinformatics, University of British Columbia, Genome Sciences Centre, 100-570 West 7th Avenue, Vancouver, British Columbia V5Z 4S6, Canada
6. US Department of Energy Joint Genome Institute, Berkeley, CA, USA 94720
7. School of Biological Sciences, University of Utah, Salt Lake City, Utah, USA
8. Department of Microbiology & Immunology, University of British Columbia, 2552-2350 Health Sciences Mall, Vancouver, British Columbia V6T 1Z3, Canada
9. Genome Science and Technology Program, University of British Columbia, 2329 West Mall, Vancouver, BC V6T 1Z4, Canada
10. ECOSCOPE Training Program, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z3
11. Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven

* These authors contributed equally to this work

$ Current affiliation: Department of Chemistry and Biochemistry, Thermal Biology Institute, and Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA 59717

Abstract

The anaerobic oxidation of methane coupled to sulfate reduction is a microbially mediated process requiring a syntrophic partnership between anaerobic methanotrophic (ANME) archaea and sulfate reducing bacteria (SRB). Based on genome taxonomy, ANME lineages are polyphyletic within the phylum *Halobacterota*, none of which have been isolated in pure culture. Here we reconstructed 28 ANME genomes from environmental metagenomes and flow sorted syntrophic consortia. Together with a reanalysis of previously published datasets, these genomes enabled a comparative analysis of all marine ANME clades. We determined the genomic features which separate ANME from their close methanogenic relatives and identify what differentiates the ANME clades. Large multiheme cytochromes and bioenergetic complexes predicted to be involved in novel electron bifurcation reactions are well-distributed and conserved in the ANME archaea. Significant variations in the anabolic C1 metabolism exist between different ANME groups. Our analysis raised the possibility that methylotrophic methanogenesis in the *Methanosarcinales* may have evolved from a methanotrophic ancestor.

Introduction

Anaerobic oxidation of methane (AOM) coupled to sulfate reduction is a key microbiological process in ocean sediments that controls the amount of methane released into overlying waters and the atmosphere. However, despite the global relevance and distribution of this process, including terrestrial habitats, there are currently no strain isolates that carry out AOM with sulfate. Thus, our understanding of the physiological and biochemical basis for AOM has advanced much more slowly than it has for many other microbially-mediated biogeochemical processes. Twenty years ago, strong evidence emerged that archaea may be involved in AOM based on stable isotope measurements of archaeal lipids and small subunit ribosomal RNA (16S rRNA) gene clone libraries from marine methane seeps (Hinrichs et al., 1999). Shortly thereafter fluorescence *in situ* hybridization from methane seep environments provided microscopic evidence for the existence of a prevalent interdomain consortia consisting of an archaeon related to known methanogens and a bacterium related to sulfate reducing bacteria (Boetius et al., 2000; Orphan et al., 2001b). These discoveries led to the current paradigm that sulfate-dependent AOM is carried out by anaerobic methanotrophic (ANME) archaea in a syntrophic partnership with sulfate reducing bacteria (SRB).

Subsequent work has expanded our understanding diversity and activities of the ANME archaea and lead to various theories pertaining to the biochemical mechanisms underlying the syntrophic interactions between ANME and SRB. Diversity surveys have suggested that ANME are polyphyletic with three distinct clades (ANME-1, 2, and 3) within the *Halobacterota*. 16S rRNA gene phylogenies support ANME-1 as a family-level clade, ANME-2 represent two distinct families within the *Methanosarcinales*, and members of ANME-3 are a novel genus closely related to *Methanococcoides* within the family *Methanosarcinaceae* (Knittel and Boetius, 2009; Orphan et al., 2001a). Initial 'omic analysis of fosmid

libraries from ANME organisms supported "reverse methanogenesis" as the biochemical model for the methane oxidation pathway in AOM (Hallam et al., 2004; Meyerdierks et al., 2010, 2005). Subsequent analysis of more complete ANME genomes from enrichment cultures (Krukenberg et al., 2018; Stokke et al., 2012; Wang et al., 2013), or metagenome assembled genomes (MAGs) from AOM habitats have added to our understanding of some of the major groups of ANME, and further refined the reverse methanogenesis hypothesis (Borrel et al., 2019; McGlynn, 2017).

The limited number of ANME genomes currently available relative to their 16S rRNA gene diversity leads to questions about whether or not the observations made in previous studies represent conserved features of the ANME archaea, or are skewed by the relatively small sample size and the incomplete or biased nature of metagenomic binning methods. In order to develop a better model for the evolution and metabolic capabilities of the ANME archaea we performed a large comparative analysis of the most complete set of ANME genomes to date, encompassing 39 reconstructed MAGs, fosmids, and amplified single aggregate genomes (ASAGs), and combine-assembled single amplified genomes (Co-SAGs), more than doubling the previously available genomic information. This analysis includes representatives from all marine ANME groups including ANME-1, 2a, 2b, 2c, and 3. Using this expanded dataset we constructed a more robust phylogenetic framework for the ANME archaea and analyzed the differences in metabolic potential that exist within and between ANME clades, as well as between ANME and their methanogenic relatives. The distribution of various mechanisms of energy conservation and potential adaptations to syntrophic associations with SRB were documented across different ANME lineages. Based on these genomic observations we present a model for ANME metabolism for the different lineages and highlight important aspects of the ANME archaea physiology that require further attention.

Results

Genome-resolved diversity of the methanotrophic archaea

We reconstructed 28 new ANME genomes from metagenomic datasets and fluorescence-activated cell sorting of single aggregates (Hatzenpichler et al., 2016), as well as from reanalyzed publicly available metagenomic data from the sequence read archive (SRA) and the MG-RAST analysis server (see Methods for details). These genomes were combined with 11 previously published marine ANME genomes recovered from diverse environments (Dombrowski et al., 2017; Krukenberg et al., 2018; Meyerdierks et al., 2010; Parks et al., 2017; Wang et al., 2013; Yu et al., 2018) to generate a set of 39 ANME genomes for comparative genomics, representing all of the major currently described clades (ANME-1, 2, 3) and subclades (e.g. ANME-2a, b, c and multiple clades in ANME-1) frequently detected in marine sediments (**Table 1)**(Knittel and Boetius, 2009). In most analyses we also include the recently described alkane-oxidizing "*Candidatus* Syntrophoarchaeum" (Laso-Pérez et al., 2016), as well as the nitrate-reducing freshwater ANME relatives known as ANME-2d or "*Candidatus* Methanoperedens" (Arshad et al., 2015; Haroon et al., 2013) and their marine relatives known as "*Candidatus* Argoarchaeum" (Chen et al., 2019; Dombrowski et al., 2017). At least one genus from each ANME subgroup was assigned a name, formal etymology can be found in **Supplementary file 1**.

The draft ANME genomes sequenced and assembled here were between 46% and 98% complete (mean: 78%, median: 80%) as determined by the presence of a set of marker proteins common to *Halobacterota*. Many of the less complete genomes originated from sequencing of individual flow sorted aggregates (**Table 1**). Most of the genomes did not contain duplicated marker genes, with the exception of the three genomes derived from sequenced fosmids, which encoded many duplicates (**Table 1**). The fosmids representing these genomes have consistent nucleotide signatures and, in many cases, contain

overlapping regions that suggests that they are derived from multiple strains rather than from a single clonal population. While genome incompleteness can impact the accurate reconstruction of ANME metabolism, we found reproducible trends in gene presence, absence, and synteny across the different ANME lineages, with multiple genomes for each ANME clade, often originating from different studies and habitats. Taxonomic assignments in **Table 1** were made consistent with GTDB release 89 using analysis of relative evolutionary divergence (RED) (Parks et al., 2018).

Members of the ANME-1 were originally described in 16S rRNA gene surveys of methane seep sediments (Hinrichs et al., 1999) and representatives of the ANME-1 were among the first to be genomically characterized (Hallam et al., 2004). ANME-1 16S rRNA genes have since been identified in marine cold seep environments, diffusive margin sulfate-methane transition zones, deep sea hydrothermal vents, and select anoxic terrestrial ecosystems (Takeuchi et al., 2011). With our updated genomic dataset there are now 19 representative genomes from ANME-1, recovered from eight locations, including deep-sea methane cold seeps, hydrocarbon-impacted hydrothermal vents and cold seeps, a mud volcano, and a hot, deep gold mine aquifer. Collectively, these genomes are highly diverse at the sequence level with the majority being at most 60% similar to each other, based on pairwise sequence similarity and non-homologous gene content between genomes (**Table 1—figure supplement 1**). These genomes represent six genera as determined by analysis of their RED values. Recently proposed nomenclature based on a single ANME-1 genome from fosmid sequences placed ANME-1 within their own order *Methanophagales* (Adam et al., 2017), and this is consistent with GTDB release 89. The 19 ANME-1 genomes analyzed here represent a single family-level division within that order. We propose to conserve the family and genus-level designation implied by "*Methanophagales*" with

the ANME-1 genome from Meyerdierks *et al.* 2010 belonging to the genus *Methanophaga* within the family *Methanophagaceae* (**Table 1**).

Marine members of the ANME-2 are within the order *Methanosarcinales* and are composed of subgroups a, b, and c, first designated in Orphan et al. 2001 based on 16S rRNA gene sequences (Orphan et al., 2001a). The ANME-2a and 2b form a family-level division with two genus-level clades corresponding to ANME-2a and 2b, recovered from six different locations including two methane cold seep sites, two submarine mud volcanoes, a hydrocarbon seep and shallow coastal sandy sediments (**Table 1**). We propose the name "*Candidatus* Methanocomedens" for the dominant genera corresponding to ANME-2a with name propagating to the family level as *Methanocomedenaceae*. Recent coupled fluorescence and electron microscopy analyses have revealed distinct ultrastructural features of the ANME-2b (McGlynn et al., 2018), and in conjunction with phylogenetic information, we further propose the genus name "*Candidatus* Methanomarinus" for ANME-2b. The six "*Ca.* Methanocomedens" genomes have AAI values which range between 70% and 97% indicating that there are multiple distinct species, whereas the three "*Ca.* Methanomarinus" genomes have >98% AAI indicating that they represent strains of the same species (**Table 1—figure supplement 1**). The ANME-2c form a separate family (84% AAI among the 8 genomes analyzed in this study) representing two genus-level groups. For the dominant ANME-2c we propose the genus name "*Candidatus* Methanogaster" with name propagating to the family level as *Methanogasteraceae*.

"*Ca.* Methanoperedens nitroreducens" is a representative of the family *Methanoperedenaceae* formerly known as ANME-2d, within the order *Methanosarcinales*. These are the only known ANME that do not couple AOM in syntrophy with partner SRB, instead coupling the oxidation of methane with the

reduction of nitrate, iron oxides or manganese oxides in freshwater environments (Cai et al., 2018; Ettwig et al., 2016; Haroon et al., 2013). The marine sister group of Methanoperedenaceae, GoM Arc I (also known as AAA), was recently described as an anaerobic ethane degrader, and contains two genera "*Candidatus* Argoarchaeum" (Chen et al., 2019; Dombrowski et al., 2017) and the thermophilic "*Candidatus* Ethanoperedens"(Hahn et al., 2020). These clades are not specifically considered in the present work as they do not appear to be marine methanotrophs and have been extensively discussed in a recent study (Leu et al., 2020).

ANME-3 is the ANME group most recently diverged from known methanogens. They are closely related by their 16S rRNA genes (94-96% similarity) to *Methanococcoides*, *Methanosalsum*, and *Methanolobus*, with 65% average AAI, indicating these organisms represent a novel genus within the family *Methanosarcinaceae*, for which we propose the name "*Candidatus* Methanovorans" (**Table 1, Table 1—figure supplement 1**). The two genomes were both recovered from the Haakon Mosby Mud Volcano where this group was originally described on the basis of 16S rRNA gene sequences (Niemann et al., 2006). This ANME genus forms consortia with bacteria from the *Desulfobulbus* group (Lösekann et al., 2007).

Phylogenetic reconstruction of 16S rRNA gene sequences, concatenated marker genes, DNA-directed RNA polymerase subunit beta (RpoB) and the methane activating enzyme McrA were constructed to demonstrate the evolutionary relationship between ANME and other related archaea (**Figure 1**). All marker gene sets show remarkably similar phylogenies, with ANME-1 forming the deepest branching group of ANME, while ANME-2 and 3 group within the order *Methanosarcinales*. Importantly, in agreement with previous studies, ANME-3 reproducibly branches well within the family

Methanosarcinaceae, in agreement with the AAI analysis described above. The McrA phylogeny deviates somewhat from this pattern, with ANME-1 falling further outside of the traditional methanogens, but grouping with some McrA from the H_2-dependent methylotrophic methanogens as has been previously reported (Borrel et al., 2019; Boyd et al., 2019; Spang et al., 2017). The McrA in all ANME groups are more similar to methanogens than the recently described divergent McrA homologs in various uncultured archaea that are thought to utilize longer chain alkane substrates (Chen et al., 2019; Dombrowski et al., 2017; Laso-Pérez et al., 2016). With this general phylogenetic framework, we set out to characterize the major conserved features of the ANME energy metabolism.

ANME energy metabolism

Biochemically it is assumed that the ANME archaea oxidize methane to CO_2 and pass electrons through an unconfirmed mechanism to their SRB partners to reduce sulfate. Metabolic reconstruction of a limited number of ANME draft genomes from environmental samples and enrichment cultures using multiple 'omics approaches have shown that they contain and express the same genes used for methanogenesis (Hallam et al., 2004; Haroon et al., 2013; Krukenberg et al., 2018; Meyerdierks et al., 2010; Pernthaler et al., 2008; Stokke et al., 2012). Detailed enzymatic studies of the methanogenic pathway have shown that all 7 steps are reversible (Holler et al., 2011; Kojima et al., 2014; Thauer, 2011), including the methane producing step catalyzed by methyl-coenzyme M reductase (Mcr) (Scheller et al., 2010), which had previously been predicted to be irreversible. These findings reinforce the 'reverse methanogenesis' hypothesis (Hallam et al., 2004; Hoehler et al., 1994; Krüger et al., 2003; Zehnder and Brock, 1980), in which ANME use the same methanogenic enzymes to oxidize, rather than produce, methane. This model offers the most likely pathway for carbon oxidation in ANME; however the mechanisms by which these archaea conserve energy from this process, and the way by which methane

is activated and methane-derived electrons are transferred to their syntrophic sulfate-reducing bacterial partners remain open questions.

It has been pointed out in recent reviews that a simple reversal of the methanogenesis pathway does not represent a viable basis for an energy metabolism in the ANME archaea, as the exact reversal of a process that results in net ATP generation must result in ATP loss (McGlynn, 2017; Timmers et al., 2017). Bioenergetic novelty beyond a wholesale reversal of methanogenesis therefore must exist. In the following discussion we break down the ANME energy metabolism into three phases and discuss the conserved features of these phases across our collection of ANME genomes (**Figure 2**). In the first phase, methane is oxidized to CO_2 and all eight electrons are deposited on cytoplasmic electron carriers in an endergonic process that requires an investment of energy. In the second phase, cytoplasmic electron carriers are re-oxidized in an exergonic process that reduces a set of intermediate electron carriers, recovering the energy invested in the first phase and conserving additional energy as an ion motive force. In the final phase, electrons must be discarded, likely not involving energy gain or loss. We use this division of energy metabolism as an organizing framework throughout this work.

Energy metabolism phase 1: The conserved C1 machinery of the methanogenesis pathway in ANME archaea

Function of the methanogenesis pathway in ANME archaea

Our analysis of this expanded set of ANME genomes is consistent with early genomic work that identified reverse methanogenesis as the most likely pathway of carbon oxidation in the ANME archaea (Hallam et al., 2004; Meyerdierks et al., 2010; Wang et al., 2013). Genes for all seven steps of the methanogenesis pathway were found in all of the ANME genomes (**Figure 3**). The only consistent exception is F_{420}-dependent methylenetetrahydromethanopterin reductase (Mer), which is absent from

all 19 ANME-1 genomes as well as the "*Ca.* Syntrophoarchaeum", as observed previously (Hallam et al., 2004; Laso-Pérez et al., 2016; Meyerdierks et al., 2010; Stokke et al., 2012). This modification of the canonical methanogenesis pathway is a common feature of the entire class Syntrophoarchaeia. Some ANME contain paralogous copies of enzymes carrying out certain steps of the pathway and these are indicated in **Figure 3**. Notably, none of the ANME genomes contain the specific methyltransferases for methanol (Sauer et al., 1997), methylamines (Paul et al., 2000) or methylated sulfur compounds (Fu and Metcalf, 2015) used for methylotrophic methanogenesis in the *Methanosarcinaceae*. This strongly argues against methylated compounds as intermediates or end products of methane oxidation in any known ANME group.

To understand whether the transition to a methanotrophic lifestyle is accompanied by a significant diversification of the enzymes catalyzing the reactions of the reverse methanogenesis pathway we built phylogenetic trees of the enzymes involved in each step. With regard to the first step, i.e. the presumed involvement of the McrA in methane activation, the McrA phylogeny largely tracks with phylogenetic marker trees, except for ANME-1 (**Figure 1**). The second step of the pathway is carried out by the N^5-methyl-H_4MPT:coenzyme M methyltransferase (Mtr) complex (**Figure 3**). All ANME groups contain Mtr homologs that are phylogenetically consistent with their genome phylogeny (**Figure 4A**). However, a highly divergent second copy of the entire Mtr complex exists that was first observed in the genome of an ANME-2a from an enrichment culture (Wang et al., 2013). We find this second "Mtr2" to be sporadically distributed through the ANME-2. Absent from the other ANME-2a genomes, Mtr2 is present in ANME-2c, one ANME-2d and both "*Candidatus* Argoarchaeum". These Mtr2 complexes form a monophyletic group that is phylogenetically distinct from all previously described methanogens yet still contains the same gene synteny in the eight-subunit cluster. In addition, all the Mtr2 gene clusters

are followed downstream of MtrH2 by a highly divergent Mer2, which catalyzes the next step in the methanogenesis pathway. All ANME-2 and 3 contain a less divergent copy of Mer that tracks with their genome phylogeny (Mer1) (**Figure 4B**).

The next fourth step of the pathway is catalyzed by the F_{420}-dependent N^5,N^{10}-methylene-H_4MPT dehydrogenase (Mtd) enzyme (**Figure 3**). Only a single copy of the gene encoding Mtd was found in any genome, and the phylogeny of the predicted protein sequence is largely congruent with the genome phylogenies (**Figure 4C**). The fifth step in the pathway is catalyzed by N^5,N^{10}-methenyl-H_4MPT cyclohydrolase (Mch). The analysis of early fosmid libraries revealed a gene on an ANME-2c-assigned fosmid encoding Mch (Meyerdierks et al., 2005), which was highly divergent from an Mch located on an ANME-2-assigned fosmid from a previous study (Hallam et al., 2004). This led the authors to question whether the Mch had diverged rapidly between ANME-2c and different ANME-2 groups, or whether multiple Mch copies were present as paralogs within ANME-2c genomes. Interestingly, ANME-2a, 2b, 2c and 2d all share a well-supported monophyletic group of Mch genes (Mch2) that are very different from those of closely related methanogens (Mch1) (**Figure 4D**). The Mch2 cluster corresponds to the gene identified as being closely related to Mch found in *Archaeoglobus* (Meyerdierks et al., 2005). The ANME-2c genomes analyzed here also contain a copy of Mch1 that is closely related to those found in the methanogenic *Methanosarcinaceae*, which corresponds to the gene found in the first set of ANME-2-assigned fosmids (Hallam et al., 2004). Apparently ANME-2c contain both copies of this gene, while the ANME-2a, 2b, and 2d only contain the divergent Mch2. ANME-3 contain a copy very similar to those of their close methanogenic relatives.

The reaction catalyzed by Mch is a curious step of the methanogenesis pathway to have a strongly supported, ANME-specific clade of enzymes. The cyclohydrolase reaction is thought to occur essentially at equilibrium (Thauer, 2011), so it is unclear what evolutionary pressure would result in such a stark difference between Mch2 enzymes in ANME-2s and their closely related methanogenic relatives. The pterin moiety of the H_4MPT analog could vary between ANME and the methanogens in the *Methanosarcinaceae*. At least five pterins are known to be found in H_4MPT analogs in methanogenic archaea: methanopterin, sarcinopterin, tatiopterin-I, tatiopterin-O and thermopterin (Raemakers-Franken et al., 1991). However, this level of sequence variation is not observed in other enzymatic steps of the pathway, which might be expected if the divergent Mch2 was the result of a significantly different form of C1-carrier in ANME.

Divergent paralogs of formylmethanofuran-tetrahydromethanopterin N-formyltransferase (Ftr) were found in ANME-1, 2a and 2b. These Ftr2 clustered together in the phylogenetic tree with Ftr genes from *Archaeoglobales* and deeper branching hydrogenotrophic methanogens such as *Methanopyrus kandleri* and *Methanothermobacter marburgensis* (**Figure 4E**). These archaea all contain both Ftr1 and Ftr2, and only the Ftr1 versions have been biochemically characterized to our knowledge (Mamat et al., 2009). In the cases where transcriptomic information is available, Ftr1 is more highly expressed in ANME, and is therefore expected to be the dominant version utilized in ANME energy metabolism under the AOM conditions tested (**Supplementary file 2**).

The seventh and final step of the methanogenesis pathway is carried out by formyl-methanofuran dehydrogenase (**Figure 3**). Two major variants of formylmethanofuran dehydrogenase are present in methanogens that contain either tungsten or molybdenum metal centers in their active sites (Fwd and

Fmd, respectively), and multiple paralogs of both can be found in methanogens such as *Methanosarcina acetivorans* (Rohlin and Gunsalus, 2010). Based on homology to versions of these enzymes in *M. acetivorans*, it appears that ANME largely contain the Fwd version, which have been detected in proteomic analyses of methane seeps (Glass et al., 2014) (**Figure 4F**). Some members of the ANME-2a, 2b, and 2c have the Fmd version as well, and in the ANME-2c the genes encoding Fmd and Fwd occur together in a single large gene cluster. As was observed in previous ANME-1 genomes, the FwdFG subunits are not present in any ANME-1 (Borrel et al., 2019; Hallam et al., 2004; Meyerdierks et al., 2010).

Based on the above observations we conclude that the transition from methanogenesis to methanotrophy required relatively little biochemical novelty within the central C1-carrying pathway of ANME energy metabolism. The loss of Mer in ANME-1 (step 3, **Figure 3**) remains the single major variation to the central C1-carrying pathway of ANME energy metabolism. Some paralogs exist, but with the exception of Mch2 in some ANME-2, these are less well conserved and less well expressed than those previously found in methanogenic archaea. While McrA in ANME-1 is slightly incongruous with its genome phylogeny, and was found to bind a modified F_{430} cofactor (Shima et al., 2012), we see little evidence for significant changes in ANME-2 or 3, suggesting these modifications are not necessary for using Mcr to activate methane. These results are broadly consistent with biochemical studies that have demonstrated the reversibility of the enzymes in this pathway (Kojima et al., 2014; Scheller et al., 2010; Thauer, 2011), and suggests that there has likely been little specialization in terms of their directionality during the evolution of the ANME archaea.

Differing roles for MetF in ANME archaea

Methylenetetrahydrofolate reductase (MetF) has been proposed to act in the third step as a replacement for Mer in ANME-1 (Stokke et al., 2012; Timmers et al., 2017) and *metF* is expressed at similar levels as other genes in the reverse methanogenesis pathway (Krukenberg et al., 2018). MetF is structurally similar to Mer and completes the same step of the Wood-Ljungdahl pathway in bacteria but uses NADPH as the electron donor rather than $F_{420}H_2$, and interacts with C1-bound tetrahydrofolate (H_4F) instead of tetrahydromethanopterin (H_4MPT).

MetF is not only found in ANME-1, but also in other ANME groups and methanogenic members of the *Methanosarcinaceae* where it is expected to function as a methylene-H_4F-interacting enzyme involved in anabolic processes (Buchenau and Thauer, 2004). Since the potential Mer/MetF switch in ANME-1 appears to be the only significant modification to the central carbon oxidation pathway in ANME we investigated the distribution, phylogenetic placement, and genomic context of MetF in ANME in order to try and better understand the evolutionary history of these proteins. MetF homologs found in ANME-1 were clearly very different at the sequence level from those found in other ANME groups (**Figure 5A**), and, interestingly, the ANME-2c universally lack MetF of either type. All ANME MetF are found next to a MetV gene which is a common feature of MetF in acetogenic bacteria, and evidence suggests a complex forms between the proteins encoded by these two genes (Mock et al., 2014).

Two phylogenetic trees were constructed with sequences of each of these MetF groups along with the most closely related homologs in the NCBI NR database (**Figure 5B,C**). The ANME-1 MetF homologs branched together with a diverse group of proteins found exclusively in uncultivated archaeal genomes, most of them identified as "*Ca.* Bathyarchaeota", "*Ca.* Verstraetearchaeota", and various members of

the "*Ca.* Asgardarchaeota". The MetF from the other ANME were all found within a monophyletic group containing other closely related *Methanosarcinaceae*. MetF and MetV in the *Methanosarcinaneae* are found in gene clusters with other H_4F-interacting enzymes that carry out important C1 reactions in biosynthesis, which led to the conclusion that H_4F is used for biosynthesis in the *Methanosarcinaceae* (Buchenau and Thauer, 2004). This clustering of anabolic C1 genes is preserved in many of the ANME-2 and 3, and we infer a similar function. For a detailed discussion of C1 anabolic metabolism in ANME, see below.

The phylogeny of MetF proteins in ANME is best explained by two separate acquisitions of these genes. MetF in ANME-2 and 3 were likely acquired in the common ancestor of these organisms and the other *Methanosarcinaceae* along with other enzymes to utilize H_4F-bound C1 moieties for the purpose of biosynthesis. Based on its distribution in uncultured archaea of different phyla, as well as the paucity of other H_4F-interacting proteins in ANME-1, there is good support for a different, possibly catabolic, function of MetF in ANME-1. Due to the structural similarity between H_4F and H_4MPT (Shima et al., 2000), it is possible that the ANME-1-type MetF has evolved to react with H_4MPT instead of H_4F. A switch between H4F and H4MPT as a carbon carrier is not unprecedented, as serine hydroxymethyl transferase (GlyA) has different versions specific to either H_4F (Buchenau and Thauer, 2004) or H_4MPT (Angelaccio et al., 2003; Hoyt et al., 1986). Additionally, MtdA in methylotrophic bacteria has been shown to react with either H_4F or H_4MPT (Vorholt et al., 1998).

Energy metabolism phase 2: Cytoplasmic electron carrier oxidation and energy conservation

After methane is oxidized to CO_2 in the pathway described above, four electrons will be found on two molecules of $F_{420}H_2$, two electrons on reduced ferredoxin (Fd^{2-}), and two electrons on the reduced forms

of coenzyme M and coenzyme B (CoM-SH, CoB-SH). If the proposed Mer/MetF switch in ANME-1 is correct, then it is possible that MetF may have produced a reduced NADPH instead of one of the $F_{420}H_2$ (**Figure 2**). Importantly, these C1 oxidation reactions do not produce energy for the cell and will in fact consume sodium motive force at the step catalyzed by Mtr. This investment of energy helps drive the reaction in the oxidative direction, and the effect of energy investment can be seen in the decreased redox potential of the eight methane-derived electrons. The CH_4/CO_2 redox couple has a standard state midpoint potential of -220mV, and the average midpoint potential of the electrons once transferred to their cytoplasmic electron carriers is approximately -340mV (two on Fd^{2-} (-500mV), four on $F_{420}H_2$ (-360mV) and two on CoM-SH/CoB-SH (-145mV)) (Thauer, 2011). The next phase of ANME energy metabolism, the re-oxidation of these reduced cytoplasmic electron carriers, must conserve sufficient energy to overcome the loss at Mtr and support growth.

This is very similar to the situation in the energy metabolism of methylotrophic methanogens in the *Methanosarcinaceae* which are close relatives of ANME-2 and 3. In these methanogens methyl groups are transferred from substrates such as methanol or methylamines onto CoM via substrate-specific methyltransferases. The methyl group is subsequently oxidized to CO_2 via a reversal of the first six steps of the methanogenesis pathway, consuming sodium motive force at Mtr and producing two $F_{420}H_2$ and one Fd^{2-} that must be re-oxidized coupled to energy conservation. In most cases, these electrons are transferred onto membrane-bound methanophenazine as an intermediate electron carrier in processes that lead to ion motive force generation. This ion motive force is then used to produce ATP via ATP synthase. The only oxidation reaction in ANME which does not normally occur in any methanogen is the production of the heterodisulfide (CoM-S-S-CoB) from the free sulfides CoM-SH and CoB-SH. Many of the strategies of energy conservation coupled to cytoplasmic electron carrier oxidation that

have been characterized in the methylotrophic methanogens appear to be conserved in the ANME archaea.

$F_{420}H_2$ oxidation

$F_{420}H_2$ oxidation for the purpose of energy conservation in methylotrophic methanogens can occur via the $F_{420}H_2$:methanophenazine oxidoreductase complex (Fpo) or the F_{420}-reducing hydrogenase (Frh) (**Figure 6A**). $F_{420}H_2$ oxidation by the cytoplasmic Frh produces H_2 which diffuses out of the cell and is subsequently oxidized on the positive side of the membrane by a membrane-bound hydrogenase (Vht) (Kulkarni et al., 2018). Our data does not support this occurring in any known ANME groups given the lack of both the NiFe hydrogenase subunit of Frh and Vht-like hydrogenases.

This leaves Fpo as the much more likely candidate for $F_{420}H_2$ oxidation in ANME. Fpo is a homolog of respiratory complex I, and in the *Methanosarcinaceae* it couples the transfer of electrons from $F_{420}H_2$ to the membrane soluble electron carrier methanophenazine with the translocation of protons across the cell membrane. In sulfate-reducing *Archaeoglobales* the homologous $F_{420}H_2$:quinone oxidoreductase complex (Fqo) utilizes a membrane-soluble quinone acceptor instead of methanophenazine (Baumer et al., 2000; Brüggemann et al., 2000). These complexes are conserved across all ANME groups (**Figure 6—figure supplement 1**), and phylogenetic analysis of Fpo shows that homologs from ANME-2a, 2b, 2c and 3 are most similar to the *Methanosarcinaceae*, suggesting they also may utilize methanophenazine. The homologs from ANME-1 were distinct from those of the other ANME clades and were most similar to the Fqo described from *Archaeoglobales* as previously reported (Hallam et al., 2004) (**Figure 6—figure supplement 2**). Consistent with this possibility these ANME-1 genomes also contained homologs of the futalosine pathway used for menaquinone biosynthesis that are absent in the other marine ANME

clades. This suggests that ANME-1 use a quinone as their membrane-soluble electron carrier as was previously suggested (Timmers et al., 2017). In either case, Fpo and Fqo are expected to be important points of $F_{420}H_2$ oxidation and membrane energization in all ANME groups.

Ferredoxin oxidation

Similar to $F_{420}H_2$ oxidation, several mechanisms are known for coupling ferredoxin oxidation to energy conservation in methanogens. Four major pathways have been proposed: 1) a hydrogen cycling mechanism using energy conserving hydrogenase (Ech) (Kulkarni et al., 2018), 2) oxidation with a modified version of the *Rhodobacter* nitrogen fixation (Rnf) complex (Li et al., 2006), 3) an Fpo-dependent pathway (Welte and Deppenmeier, 2011a), and 4) a heterodisulfide reductase (Hdr)-mediated electron confurcation (Yan and Ferry, 2018)(**Figure 6B**). The Ech model is easily ruled out by the lack of Ech homologs in all marine ANME genomes, although these are present in the freshwater ANME-2d genomes (Arshad et al., 2015; Haroon et al., 2013).

Genomes from many members of the ANME-2a, 2b, 2c and 3 contained homologs of the Rnf complex (**Figure 6—figure supplement 1**). Rnf was first characterized in bacteria as an enzyme that performed the endergonic transfer of electrons from NADH to ferredoxin by dissipating sodium motive force (Biegel et al., 2011). In methanogens however Rnf is believed couple the exergonic electron transfer from ferredoxin to methanophenazine with the endergonic translocation of sodium ions to the positive side of the cytoplasmic membrane. This activity in methanogens involves of a multiheme cytochrome *c* subunit which is present in their Rnf gene clusters (Li et al., 2006), along with a small conserved membrane integral protein (MA0665 in *M. acetivorans*). This enzyme complex operates in methylotrophic members of the *Methanosarcinaceae* that do not conduct hydrogen cycling, and in some cases is required

for energy conservation during acetoclastic methanogenesis (Schlegel et al., 2012; Welte and Deppenmeier, 2014).

Rnf was not found in the ANME-1 genomes analyzed here, with the exception of two genomes from a genus-level subclade recovered from a South African gold mine aquifer (SA) and from a hydrocarbon seep from the Gulf of Mexico (GoMg4) (**Table 1**; **Figure 6—figure supplement 1**). Rnf gene clusters in ANME-2a, 2b and 3 contain homologs of the cytochrome *c* subunit and MA0665 normally found in methanogens, but surprisingly both were missing in the Rnf gene clusters from all ANME-2c and the two ANME-1. Homologs for these subunits were not identified elsewhere in any of these genomes. Based on the current information, it is unclear what reaction the ANME-2c and ANME-1 Rnf are performing since it has been demonstrated that this cytochrome *c* is involved in the electron transfer to methanophenazine (Wang et al., 2011). It is possible that Rnf retains the ability to transfer electrons from ferredoxin to a membrane bound electron carrier in these lineages, or alternatively they could function as ferredoxin:NAD^+ oxidoreductase, as found in bacteria. It is unclear what role the latter function could play in our current model of ANME metabolism.

An additional ANME-specific modification to the *rnf* gene cluster is the inclusion of a membrane-bound *b*-type cytochrome in ANME-2a, 2b, and 3. This type of cytochrome is generally involved in electron transfer between membrane-bound and soluble electron carriers, and has no closely related homologs in methanogens. If the protein encoded by this gene is incorporated into the Rnf complex, it could be of great importance to the flow of electrons in these groups. This observation is particularly striking in ANME-3 since their close methanogenic relatives in the Methanosarcinaceae lack this *b*-type cytochrome. This means that ANME-3 has acquired this subunit by horizontal gene transfer from an

ANME-2a or 2b, or that all of the related methanogens have lost it. Whichever evolutionary scenario is correct, this represents an important ANME-specific modification to a key bioenergetic complex.

Because the majority of ANME-1 do not contain an Rnf complex an alternative mechanism is needed to explain how ferredoxin is recycled. One option is the Fpo-dependent mechanism proposed in *Methanosaeta thermophila* where ferredoxin is thought to be oxidized using an Fpo complex without the FpoF subunit (Welte and Deppenmeier, 2011a). If this pathway operates in ANME-1 the reduced ferredoxin would donate electrons directly to the iron sulfur clusters found in the FqoBCDI subunits. Since FqoF homologs are encoded in the ANME-1 genomes, and are expected to be used in complete Fqo complexes to oxidize $F_{420}H_2$ as described above, this Fd^{2-} oxidation strategy would require some Fqo complexes to have FqoF bound, while others do not.

Another possibility would be for soluble FqoF to act as a Fd^{2-}:F_{420} oxidoreductase, and then the subsequent oxidation of $F_{420}H_2$ by a normal Fqo complex (**Figure 6B**). This pathway has been proposed in *Methanosarcina mazei* based on the Fd^{2-}:F_{420} oxidoreductase activity of soluble FpoF (Welte and Deppenmeier, 2011b). In either of these Fpo/Fqo-dependent Fd^{2-} oxidation pathways, the oxidation of Fd^{2-} would result in the reduction of a membrane-bound electron carrier coupled to energy conservation in the form of proton translocation by the Fpo/Fqo complexes.

A final possible mechanism of Fd^{2-} oxidation is through soluble heterodisulfide reductase (Hdr)-mediated flavin-based electron confurcation. This process has some biochemical support for operating in this direction (Yan and Ferry, 2018), but would constitute a reversal of the well accepted electron bifurcation mechanism used by many hydrogenotrophic methanogens. In the best characterized

examples of this process, an enzyme complex of soluble Hdr and a hydrogenase (hdrABC-mvhADG) reduce ferredoxin and heterodisulfide with four electrons sourced from two hydrogen molecules (Buckel and Thauer, 2013). If reversed, this reaction could potentially oxidize Fd^{2-}. As this mechanism would also involve CoM-SH/CoB-SH oxidation it is discussed in detail in the following section.

CoM-SH/CoB-SH oxidation

The last oxidation that needs to occur during this part of ANME energy metabolism is CoM-SH/CoB-SH oxidation to the heterodisulfide (CoM-S-S-CoB). This oxidation is the most energetically challenging because the relatively high midpoint potential of the heterodisulfide (E_0' = -145mV) rules out most methanogenic electron carriers as acceptors without the input of energy to force electrons "uphill" to a lower redox potential. It also represents the second net reaction in the ANME energy metabolism that does not occur during canonical methanogenesis (the first step being methane activation through Mcr). In methylotrophic methanogens heterodisulfide is the terminal electron acceptor for the six electrons that have passed through $F_{420}H_2$ and Fd^{2-} discussed in the previous sections.

Two enzyme systems from methanogens could potentially produce heterodisulfide by running in reverse: soluble HdrABC using a confurcation mechanism as mentioned above, or HdrDE, a membrane-bound system which would deposit electrons on methanophenazine (**Figure 6C**). ANME-2a, 2b, 2c, 2d, and 3 genomes contained HdrDE genes similar to ones from closely related methanogens. The reaction carried out by this complex in methanogens is expected to be readily reversible, with electrons from CoM-SH/CoB-SH oxidation being deposited on methanophenazine. This electron transfer reaction is slightly endergonic at standard state due to the redox potential of methanophenazine (E_0' = -165mV) being lower than that of heterodisulfide by 20mV. The slightly endergonic nature of the

electron transfer could be overcome by a quinol loop-like mechanism, where two protons are released in the cytoplasm from CoM-SH/CoB-SH oxidation and two protons are consumed on the outer face of the cytoplasmic membrane to form reduced MpH_2, thus dissipating proton motive force (Simon et al., 2008). Electron transfer from the cytoplasm to the outer face of the membrane occurs through the *b*-type cytochromes in the HdrE subunit. The HdrDE complexes are well expressed in a number of different ANME groups (**Supplementary file 2**).

HdrDE gene clusters are absent in all ANME-1 genomes. This, combined with the lack of Rnf and Ech, suggests that ANME-1 in particular could rely on HdrABC complexes for both CoM-SH/CoB-SH and ferredoxin oxidation. Since its discovery just over ten years ago (Li et al., 2008), flavin-based electron bifurcation and confurcation have been shown to be a key energy conversion point in many anaerobic metabolisms, with a great diversity of different electron donors and acceptors (Buckel and Thauer, 2018). In the complete HdrABC-MvhADG complex from methanogens the endergonic reduction of Fd (E_0' = -500mV) with H_2 (E_0' = -420mV) is driven by the exergonic reduction of heterodisulfide (E_0' = -145mV) with H_2.

In most cases a strict reversal of H_2 electron bifurcation is not possible in ANME due to a lack of MvhAG genes that encode the NiFe hydrogenase large and small subunits. The exceptions are two small subclades within ANME-2c and ANME-1 that contain NiFe hydrogenases similar to Mvh. These two subclades contain genomes recovered from different environments; a South African gold mine (SA) and the Gulf of Mexico (GoMg4) for ANME-1, versus a hydrocarbon seep off of Santa Barbara, CA (COP2) and a mud volcano from the Mediterranean (AMVER4-31) for ANME-2c. The majority of ANME-2c and ANME-1, however, lack these genes, and they are completely absent in ANME-2a, 2b and 3,

suggesting confurcation of electrons from Fd^{2-} and CoM-SH/CoB-SH to H_2 is not a dominant process in most ANME lineages.

Novel gene clusters encoding electron bifurcation/confurcation complexes

While a strict reversal of hydrogen-dependent electron bifurcation seems unlikely in ANME, the potential involvement of alternative bifurcation/confurcation reactions in ANME metabolism is supported by the broad distribution and conservation of unusual HdrABC homologs, even in those ANME genomes containing HdrDE. Alternative bifurcation reactions have been proposed in methanogens, where formate or $F_{420}H_2$ can serve as the electron donor in place of hydrogen. In the case of formate, MvhA and G are replaced by the $F_{420}H_2$-dependent formate dehydrogenase genes FdhAB (Costa et al., 2013, 2010). In the case of $F_{420}H_2$, certain HdrABC complexes might be able to interact with $F_{420}H_2$ without any additional protein subunits (Yan and Ferry, 2018).

A comparison of ANME HdrA homologs to crystal structures of the entire HdrABC-MvhADG complex purified from *Methanothermococcus wolfeii* (Wagner et al., 2017), reveal some stark differences in how these complexes may facilitate electron bifurcation. HdrA in methanogens normally consists of four domains, an N-terminal domain with an iron sulfur cluster, a thioredoxin reductase domain which binds the bifurcating FAD, and two ferredoxin domains (**Figure 7**). The MvhAG hydrogenase feeds electrons through MvhD and the C-terminal ferredoxin domain to the FAD where they are bifurcated, one pair passing up through HdrBC onto CoM-S-S-CoB, and the other through the inserted ferredoxin domain to a free soluble ferredoxin. Interestingly, the heterohexameric HdrABC-MvhADG forms dimers, and one of the cysteine ligands for the N-terminal iron sulfur cluster comes from the

thioredoxin-reductase domain of the other copy of HdrA (Cys197), indicating an obligate dimeric nature of the complex in *M. wolfeii* (**Figure 7C**).

By creating alignments of the ANME HdrA paralogs and comparing the presence of domains, conserved residues, sequence similarity, and their genomic context, we clustered the dominant HdrA homologs into 13 groups (**Figure 8A, Figure 8—figure supplement 1**). Most methanogen genomes contain 1-2 copies of *hdrABC* gene clusters, but ANME, and in particular ANME-1, have a greater abundance of *hdrA* homologs that exceed the number of *hdrBC* homologs (**Figure 8—figure supplement 2**). Some of these HdrA homologs exceed 1000 amino acids in length. In comparison, HdrA from methanogens is usually 650 amino acids in length, with some slightly larger homologs occurring due to the fusion of *hdrA* and *mvhD* (Stojanowic et al., 2003). The gene clusters containing HdrA homologs often contained HdrBC and MvhD as expected, but more unexpected was the co-occurrence of multiple copies of HdrA and the occasional presence of HdrD-like proteins (**Figures 8A, Figure 8—figure supplement 3**). Although HdrD is a fusion of HdrB and HdrC, it is not common to find the fused form in gene clusters with HdrA in methanogens. Distant homologs of the F_{420}-dependent formate dehydrogenase FdhAB were also found in HdrA gene clusters in all ANME groups. The gene cluster containing HdrA 2 and 3 along with FdhAB-like genes in ANME-2a was discussed in detail in one of the earliest studies of ANME fosmid libraries (Meyerdierks et al., 2005), and the significance of this cluster has now been substantiated by its broad conservation as well as reasonably high transcription levels (**Supplementary file 2**).

Many HdrA homologs in ANME were lacking various domains present in the crystalized complex from *M. wolfeii* (**Figure 8B**). An interesting pattern emerges in gene clusters with tandem HdrA genes (i.e.

HdrA 5 and 6, or 12 and 13), in which one copy has the N-terminal domain, but lack the Cys197 ligand for the N-terminal FeS cluster in its thioredoxin reductase domain, while the second copy lacks the N-terminal domain but has Cys197. It seems likely that these tandem HdrA genes form heterodimers, breaking the rotational two-fold symmetry found in the crystal structure (**Figure 8B**). This pattern of Cys197/N-terminal domain complementarity can also be seen in largest HdrA homologs 3 and 4, which are fusions of two HdrA genes. In both cases the N-terminal HdrA contains the N-terminal domain, but lacks Cys197, while the C-terminal HdrA contains Cys197 and no N-terminal domain, suggesting these two HdrA domains within a single continuous peptide may act as their own heterodimeric partners. Evidence for symmetry breaking in HdrABC complexes has recently been demonstrated in *Methanococcus maripaludis,* where a hydrogenase and a formate dehydrogenase can be simultaneously bound to dimerized HdrABC (Milton et al., 2018). This appears to be the result of HdrABC's ability to bind either hydrogenase or formate dehydrogenase in these methanogens, which results in a mixture of subunits, whereas in the ANME case this heteromeric form may be imposed by the heterodimer of two HdrA paralogs.

Another remarkable HdrA modification is a 500 amino acid insertion in the ferredoxin domain of HdrA 9, 10, and 11, present only in some ANME-1 genomes (**Figure 8**). This insertion shares 45% amino acid sequence identity to the NfnB subunit of the bifurcating NADH-dependent reduced ferredoxin:NADP oxidoreductase crystalized from *Thermotoga maritima* (Demmer et al., 2015). NfnB binds the b-FAD cofactor thought to be responsible for bifurcation in the NfnAB complex, potentially giving these HdrA-NfnB genes two electron bifurcation sites, and suggests NADPH as a possible electron carrier in these compounds.

Based on the extensive conservation of FdhAB subunits in Hdr gene clusters in nearly all ANME groups it is tempting to speculate that these enzymes may have an important role in CoM-SH/CoB-SH oxidation through electron confurcation. A reversal of electron bifurcation from formate would be a reasonable reaction to expect in ANME metabolism, as formate is a common syntrophic intermediate. Based on its concentration one could envision its midpoint potential to lie appropriately between that of Fd^{2-} and CoM-SH/CoB-SH to receive electrons from both. FdhA belongs to the molybdopterin oxidoreductase family, members of which act on many different substrates including formate, nitrate, and DMSO among others in assimilatory and dissimilatory processes (Hille et al., 2014). The FdhA homologs from ANME genomes are often annotated as formate dehydrogenases and retain a conserved cysteine ligand for binding the molybdenum atom (sometimes a selenocysteine in other organisms). However, they are very distantly related to any biochemically characterized enzymes, making it difficult to assign their substrate with any level of certainty. A very closely related gene cluster with HdrA and FdhAB can be found in the genomes of the methylotrophic methanogen genus *Methanolobus*, potentially providing a good opportunity to study the substrate specificity of this specific group of molybdopterin oxidoreductases in a pure culture organism. These are the only methanogens that encode this type of Hdr gene cluster, and it is worth noting in this context that the *Methanolobus* are incapable of using formate as a methanogenic substrate (Kendall and Boone, 2006).

Another possibility is confurcation of Fd^{2-} and CoM-SH/CoB-SH electrons to $F_{420}H_2$. FdhB and homologous proteins carry out oxidoreductase reactions with F_{420} (Mills et al., 2013) in F_{420}-dependent hydrogenases (FrhB), formate dehydrogenases (FdhB), F_{420}-dependant sulfite reductase (Fsr) and the Fpo complex (FpoF). The observation of FdhB homologs in gene clusters with HdrA in ANME-2d lead to the suggestion that these complexes could confurcate Fd^{2-} and CoM-SH/CoB-SH to reduce two

molecules of F_{420}, although this model proposes no role for the FdhA homologs present in the gene clusters (Arshad et al., 2015). Some HdrA gene clusters only contain FdhB and not an FdhA homolog, however these are found only in ANME-1 genomes (**Figure 8A, Figure 8—figure supplement 3**).Additionally, there is evidence that certain HdrABC complexes can interact with F_{420} without any additional subunits (Yan et al., 2017). If such a reaction were to produce $F_{420}H_2$ in ANME, its re-oxidation through the Fpo complex could be coupled to energy conservation as has been previously proposed in *M. acetivorans* (Buan and Metcalf, 2010).

We can only speculate on the functions of all these modifications, but it is clear that ANME have an exceptionally diverse potential of bifurcation/confurcation complexes at their disposal, and that the flow of electrons through these complexes will necessarily be different than in the traditional HdrABC-MvhADG complex due to domain gain and loss within the HdrA homologs, as well as the replacement of hydrogenase subunits with various other input modules. While any electron confurcation schemes through HdrABC is speculative at this point, they seem to be viable option for Fd^{2-} and CoM-SH/CoB-SH oxidation in ANME due to their widespread conservation across all ANME groups, particularly in ANME-1 that lack HdrDE and most known Fd^{2-} oxidation systems. Additionally, an examination of previously reported transcriptomic information indicates that these complexes are often expressed to levels on par with other components of the reverse methanogenesis pathway (**Supplementary file 2**). A key question that remains is what substrate the FdhA homologs in ANME act upon. A small molecule such as formate was previously proposed to serve as one mechanism of electron transfer between ANME and SRB, but no evidence of this was found by in vitro enrichment experiments (Nauhaus et al., 2005, 2002; Wegener et al., 2016).

Energy metabolism phase 3: Genomic evidence for mechanisms of syntrophic electron transfer

The most enigmatic phase of marine ANME metabolism is the interspecies electron transfer to the sulfate-reducing partner; a process which appears to necessitate the formation of conspicuous multicellular aggregates of the two organisms (Knittel and Boetius, 2009; Orphan et al., 2002). The cytoplasmic electron carrier oxidation described in the previous section will result in 8 electrons on a combination of membrane-bound MpH_2 or QH_2 and possibly some soluble electron carrier formed through an electron confurcation reaction oxidizing Fd^{2-} and CoM-SH/CoB-SH. Based on energetic considerations and precedent from other known syntrophies, the hypothesized mechanisms for the AOM syntrophy have included the diffusive exchange of small molecules such as hydrogen, formate, acetate, methanol, methylamine (Nauhaus et al., 2005), methyl-sulfides (Moran et al., 2008), zero-valent sulfur (Milucka et al., 2012), and direct electron transfer using multiheme cytochrome *c* proteins (McGlynn et al., 2015; Meyerdierks et al., 2010; Wegener et al., 2015). We assessed the genomic potential for each of these syntrophic electron transfer strategies across our expanded sampling of ANME clades.

Hydrogen transfer

Hydrogen transfer is one of the most common forms of syntrophic electron transfer (Schink and Stams, 2006). A classic mode of syntrophic growth involves hydrogenotrophic methanogens consuming H_2 produced by the fermentative metabolism of a syntrophic partner. A direct reversal of the methanogenic side of this syntrophy is not possible as the majority of ANME genomes lack any identifiable hydrogenases with the notable exception of MvhA homologs in a small set of ANME-1 and 2c genomes as mentioned in the preceding section (**Figure 8—figure supplement 3**). The first report of an ANME-1 genome from fosmid libraries contained a gene that appeared to encode an FeFe hydrogenase

(Meyerdierks et al., 2010); however, homologs of this gene were not found in any of the other ANME-1 genomes analyzed here.

This lack of hydrogenases is consistent with the lack of energy-conserving hydrogenases in the genomes of their syntrophic sulfate-reducing bacterial partners from cold seeps (Skennerton et al., 2017) and previous experimental results showing that the addition of excess hydrogen does not inhibit AOM (Meulepas et al., 2010; Nauhaus et al., 2002). Hydrogen has been shown to stimulate sulfate reduction in AOM sediments suggesting that at least some portion of the sulfate-reducing community can utilize this electron donor (Meulepas et al., 2010; Nauhaus et al., 2005; Wegener et al., 2016). In the case of the syntrophic thermophilic ANME-1-"*Candidatus* Desulfofervidus auxilii" consortium, hydrogen amendment suppresses growth of ANME-1, because the "*Ca.* D. auxilii" can grow alone on hydrogen and stops investing in electron transferring structures (Krukenberg et al., 2016). Together, genomic and physiological data suggests that the vast majority of ANME are incapable of producing hydrogen as an electron shuttle.

Formate transfer

Formate is another possible small molecule intermediate commonly involved in syntrophic electron transfer (Schink, 1997), and was predicted to be a possible intermediate in the ANME-SRB syntrophy based on early modeling studies (Sørensen et al., 2001). As described above FdhAB-like genes in association with Hdr-containing gene clusters are broadly distributed in ANME and could potentially be used to produce formate from Fd^{2-} and CoM-SH/CoB-SH and are relatively highly expressed (**Supplementary file 2**) but their substrate is currently unknown. In addition to these complexes, five of the ANME-1 genomes contained a membrane-bound formate dehydrogenase with signal sequences

predicted to mediate their secretion to the positive side of the cytoplasmic membrane. One of the ANME-1 genomes containing this formate dehydrogenase is based on early fosmid libraries that were analyzed extensively, and it was suggested to play a role in electron transfer, depicted as either being free-floating in the pseudoperiplasm, or anchored to the membrane via a *c*-type cytochrome (Meyerdierks et al., 2010). Interestingly, the gene cluster encoding these formate dehydrogenases in four other ANME-1 genomes also contain DmsC-like genes that traditionally serve as membrane anchor subunits for members of the molybdopterin oxidoreductase family. The presence of these genes in the other ANME-1 genomes makes it likely that this is how the complex is anchored and interacts with the ANME-1 membrane-bound quinol, since this is a fairly common pattern for members of this family of oxidoreductases (Grimaldi et al., 2013). Interestingly these formate dehydrogenase are distributed very sporadically through the ANME-1 instead of being confined to a specific subgroup (**Figure 8—figure supplement 3**).

This genomic evidence from different subclades seems to indicate a more plausible role for formate than hydrogen in ANME metabolism. However, like hydrogen, formate addition was not shown to inhibit AOM (Meulepas et al., 2010; Nauhaus et al., 2005, 2002) which would be expected if it was the major electron donor to the sulfate reducing partner in those sediments. The addition of formate to AOM sediments and enrichment cultures of ANME-SRB consortia has also had mixed results in terms of stimulating sulfate reduction (Meulepas et al., 2010; Nauhaus et al., 2005; Wegener et al., 2016). Additionally, genes encoding FdhC, a member of the formate-nitrate transporter family thought to be responsible for formate transport in many formate-utilizing methanogens (Wang et al., 2009; Wood et al., 2003) was absent from all ANME genomes.

Other soluble electron carriers

Acetate has been proposed as a syntrophic intermediate in sulfate-coupled AOM (Meyerdierks et al., 2010; Valentine and Reeburgh, 2000), and one could easily envision an energy conservation strategy in ANME archaea whereby methane and CO_2 are combined in a reversal of acetoclastic methanogenesis and energy is conserved via substrate-level phosphorylation through Pta and AckA as in acetogenic bacteria. Acetate thus produced could be used by the SRB partner for sulfate reduction. Two problems exist with this model that make it unlikely in any of the ANME groups represented here. First, Pta and AckA were absent from all ANME genomes. Second, this model requires that the two electrons transferred from methane to heterodisulfide (E_0' = -145mV) would be used for the reduction of CO_2 which requires ferredoxin (E_0' = -500mV). This would mean the single ATP produced by substrate level phosphorylation would have to overcome the sodium motive energy lost in the Mtr step, force electrons energetically uphill by ~350mV from the free disulfides onto ferredoxin, and have energy left over for growth. Experimental addition of acetate to AOM enrichments has not been found to decouple the partners and inhibit methane oxidation, as would be expected for a soluble extracellular intermediate (Meulepas et al., 2010; Nauhaus et al., 2005, 2002).

Methylsulfides were suggested to be a potential intermediate in AOM (Moran et al., 2008), but as noted above the methyltransferase genes for methylsulfides, methanol, and methylamines were not recovered in any ANME genomes, and experiments have not found methylsulfide-stimulated sulfate reduction in ANME-SRB enrichments (Meulepas et al., 2010). Zero-valent sulfur has also been suggested as an intermediate in marine AOM using an ANME-2 enrichment (Milucka et al., 2012). However, in vitro studies with other ANME consortia did not show a similar response (Wegener et al., 2016, 2015) and recent investigations into sulfur utilizing genes in ANME found no evidence for dissimilatory sulfate

reduction (Yu et al., 2018). If this process occurs, it is in a restricted group of ANME through a novel biochemical mechanism. Electron transfer is also possible via soluble organic shuttles. *Shewanella oneidensis* MR-1 was determined to utilize small molecule shuttles derived from menaquinone (Newman and Kolter, 2000) or flavins (Brutinel and Gralnick, 2012) that help electrons pass through the external environment to their ultimate electron acceptor. The mechanism of producing these compounds is very poorly understood. Only recently a major menaquinone-like shuttle was identified and its biosynthesis elucidated (Mevers et al., 2019). Such a shuttle-based electron transfer strategy could be potentially used for accepting electrons from the ANME Hdr complexes, but predicting the occurrence of these shuttling mechanisms from genomic information does not seem possibly with our current understanding of these processes.

Direct interspecies electron transfer

If soluble electron carriers were produced by ANME, they could simply diffuse to the SRB without any additional specialized protein systems. Specific transporters or permeases would only be needed if the compounds produced in the ANME cytoplasm are not membrane permeable. On the other hand, electrons on MpH_2/QH_2 would require an oxidation system with no analog in methanogens. In cytochrome-containing methanogens the eventual electron acceptor for these membrane-bound electrons is the heterodisulfide located in the cytoplasm, whereas the terminal electron acceptor of marine ANME would be their partner SRB located outside of the cell.

This is a respiratory challenge that is similar to the one facing bacteria that carry out extracellular electron transfer (EET) to utilize insoluble metal oxides or electrodes as terminal electron acceptors. Instead of having a soluble terminal electron acceptor that can diffuse to the cytoplasmic membrane and be reduced

by a terminal oxidase, these organisms make conduits for their electron transport chain to extend through the periplasm and outer membranes to interact directly with an electron acceptor in the extracellular space (Richardson et al., 2012).

Various systems have been discovered in gram negative bacteria that perform EET. These generally consist of a quinol:cytochrome *c* oxidoreductase enzyme, small soluble cytochrome *c* intermediates between the inner and outer membranes, and a beta barrel/decaheme cytochrome *c* protein complex that is embedded in the outer membrane (**Figure 9A**)(Shi et al., 2016). In bacteria that form large conductive biofilms, the extracellular space is also densely packed with secreted *c*-type cytochromes which can form large conductive complexes (Wang et al., 2019) . Based on the similarity of the metabolic challenge facing ANME and EET-capable bacteria, as well as the discovery of abundant analogs of bacterial EET systems in ANME genomes (McGlynn et al., 2015; Meyerdierks et al., 2010; Wegener et al., 2015), an attractive hypothesis is that the ANME-SRB syntrophy is based on direct electron transfer between the syntrophic partners. However, in contrast to other EET-capable bacteria that use metals as terminal electron acceptors, the highly specific interspecies interaction observed in the AOM consortia may require the partner SRB to contain a specific electron-transferring systems.

Potential methanophenazine:cytochrome c *oxidoreductase complexes*

In bacteria capable of EET, the quinol:cytochrome *c* oxidoreduction step is carried out by a diverse group of non-homologous enzymes. These can be as simple as CymA in *Shewanella oneidensis* MR-1, which is a four heme-binding cytochrome *c* protein with a single transmembrane helix (Schwalb et al., 2003), or more complex, involving multiple subunits such as the recently described NetBCD system in *Aeromonas hydrophila* (Conley et al., 2018). These complexes are relatively easy to replace, as CymA

knockouts in *S. oneidensis* MR-1 can be rescued by suppressor mutants that turn on the expression of completely unrelated quinol oxidoreductase complexes (Cordova et al., 2011). In *Geobacter sulfurreducens* at least two quinol:cytochrome *c* systems coexist, ImcH and CbcL, that appear to be tuned to the redox potential of different terminal electron acceptors (Levar et al., 2014; Zacharoff et al., 2016). The only common features in these systems is the presence of periplasmic cytochrome *c* or FeS binding proteins associated with, or fused to, membrane anchor subunits that facilitate electron transfer from the membrane bound quinol to the periplasmic acceptor (**Figure 9A**).

Due to the non-homologous nature of these electron transport systems in bacteria, we examined the ANME genomes for genes and gene clusters with a potential for analogous function, but adapted to the specifics of archaeal cell biology (**Figure 9B**). ANME-2a, 2b, and 2c were found to encode a membrane-bound cytochrome *b* that is closely related to the membrane-bound subunit of the Vht hydrogenase (VhtC), which in methanogens mediates the reduction of methanophenazine with electrons from H_2 (Thauer et al., 2010) (**Figure 9—figure supplement 1**). In the genomes of ANME-2a and 2b, this cytochrome *b* is followed by a multiheme *c*-type cytochrome (MHC) containing between 7 and 11 heme-binding motifs (CxxCH), instead of the *vhtAG* genes encoding the hydrogenase subunits found in *Methanosarcina*. In ANME-2c a closely related homolog for this MHC protein was found elsewhere in the genome. This conspicuous gene clustering is not found in any methanogenic archaea, and the importance of this system is supported by the high expression of the cytochrome *b* subunit reported in both ANME-2a and ANME-2c (Krukenberg et al., 2018; Wang et al., 2013). Phylogenetic analysis shows these VhtC homologs to be a closely related sister group to those found in *Methanosarcina* (**Figure 9—figure supplement 2**). These gene clusters are the clearest examples of biological novelty well conserved

in ANME that could explain the evolution of electron transfer capabilities, and we refer to them here as methanophenazine-cytochrome c oxidoreductase (Mco).

Notably, neither ANME-3 nor ANME-1 contained homologs to Mco, so this does not appear to be a universal mechanism of electron transport in all ANME. ANME-1 lacked any identifiable cytochrome *b*, while the only ones apparent in ANME-3 were multiple copies of *hdrE* and the gene associated with Rnf gene cluster mentioned above. These Rnf-associated cytochrome *b* in ANME-2a, 2b, and 3 are not found in any related methanogens. It is possible that these is also somehow involved in the oxidation of MpH_2 and in the transfer of these electrons onto the Rnf-associated cytochrome *c* for the purpose of EET. Two recent studies on *M. acetivorans* have implicated the Rnf-associated cytochrome *c* in electron transfer to Fe(III) or the artificial electron acceptor AQDS (Holmes et al., 2019; Prakash et al., 2019). Such a process may be occurring in the ANME-SRB syntrophy via a similar mechanism and explain previous results of marine ANME utilizing AQDS, Fe(III), and Mn(IV) (Beal et al., 2009; Scheller et al., 2016).

Another possible MpH_2/QH_2 oxidizing systems found in ANME genomes are diverse homologs related to the membrane-bound periplasmic nitrate reductase subunit NapH which are sporadically distributed throughout ANME genomes. NapH typically contains four transmembrane helices and two conserved cytoplasmic FeS clusters, and mediates electron flow from a menaquinol to the periplasmic nitrate reductase NapAB (Sparacino-Watkins et al., 2014). The NapH homologs in ANME are not found in suggestive gene clusters except in one of the ANME-2c genomes in which NapH was followed immediately by an 8 heme MHC, reminiscent of quinol:cytochrome *c* oxidoreductase gene clusters described in bacteria (**Figure 9B**). NapH homologs are found in both ANME-3, as well as two ANME-

1, and are worthy of investigating further due to the lack of other obvious candidates for menaquinol/methanophenazine oxidation in those genomes. However, under standard laboratory AOM conditions (Krukenberg et al., 2018), the NapH homolog found in ANME-2c E20 had very low expression levels (**Supplementary file 2**). Due to their uneven distribution in ANME genomes these NapH homologs may have other non-essential functions.

While no other gene clusters could be identified that seemed likely to carry out this MpH_2/QH_2:cytochrome *c* oxidoreduction step, the example of CymA in *S. oneidensis* MR-1 highlights how simple these systems can be while still being vitally important respiratory proteins, and how easily one system can be functionally complemented by a completely unrelated one (Cordova et al., 2011; Schwalb et al., 2003). Numerous small multiheme cytochromes are encoded and expressed in ANME genomes that either have single membrane anchors on their C-terminus or PGF-CTERM archaeosortase motifs that are predicted to covalently link them to membrane lipids. These small MHC proteins could potentially act in a similar fashion to CymA. However, without an associated or fused large membrane anchor that is homologous to previously characterized systems it is difficult to implicate them directly in membrane bound electron carrier oxidation through genomic evidence alone.

Multiheme cytochrome c *protein abundance and expression*

ANME genomes contained many (5-8 times) more genes encoding MHCs than any of their methanogenic relatives (**Figure 9—figure supplement 3**). Only the "*Ca.* Syntrophoarchaeum" and some members of the *Archaeoglobales* that are known to conduct EET contain a similar number of MHC genes among the archaea. Many homologous groups of these MHCs are only present in ANME and these other archaea. All MHC proteins are predicted to reside in the extracellular space as heme

attachment to the cytochrome *c* apoprotein occurs on the positive side of the cytoplasmic membrane (Sanders et al., 2010). Representatives of these MHCs are among the highest expressed proteins in multiple ANME lineages. Small MHC in ANME-2c E20, ANME-1 GB37 and ANME-1 GB60 were the 14th, 10th, and 18th highest expressed proteins, respectively, and contain 8, 4 and 4 heme binding motifs, respectively (**Supplementary file 2**). One specific group of small MHC proteins has widespread distribution in the ANME, "*Ca.* Syntrophoarchaeum", and some members of the *Archaeoglobales*, with no relatives in methanogenic archaea. Multiple paralogs exist in most ANME groups and these include some of the highest expressed proteins in ANME, many exceeding all methanogenesis pathway genes except Mcr (**Figure 9—figure supplement 4, Supplementary file 2**).

As with the quinol oxidation step, the intermediate small soluble cytochromes *c* vary greatly between different EET-capable bacteria, so any number of these ANME cytochromes could be capable of carrying out this intermediate step between the cytoplasmic membrane and outermost layer of the cells. A small six heme-binding cytochrome *c* protein OmcS has recently been shown to be the monomer that makes up conductive extracellular "nanowires" in *G. sulfurreducens* thought to imbue biofilms with conductive properties (**Figure 9**) (Filman et al., 2019; Wang et al., 2019). No close homolog of OmcS were identified in ANME genomes, although any number of these relatively small cytochromes could be carrying out a similar function in the extracellular space between ANME and SRB, consistent with ultrastructural observations made with heme-reactive staining (Krukenberg et al., 2018; McGlynn et al., 2015).

An 8-heme MHC in *G. sulfurreducens* known as OmcZ was shown to be required for optimal anodic current in biofilms grown on electrodes, and is secreted in biofilms and in culture supernatant (Inoue et

al., 2010). Very closely related homologs of OmcZ were found in multiple ANME-2a, 2b, and 2c genomes (**Figure 10**). The sequence similarity shared between OmcZ and the proteins found in these ANME groups is not limited to CxxCH motifs, but also extensive N- and C-terminal regions with many completely conserved residues. This level of sequence conservation is quite remarkable for homologs found in archaea and bacteria, and suggest an important conserved function of these non-heme binding domains, as well as a relatively recent inter-domain horizontal transfer. The OmcZ homolog in ANME-2c E20 is the MHC that is the 14th highest expressed protein in cultures grown under standard laboratory AOM conditions (**Supplementary file 2**). In ANME-2a and 2b the genes encoding these OmcZ homologs are found next to the enormous MHCs described below.

S-layer conduits

ANME-2a, 2b, 2c, 2d, and 3 genomes also encode for exceptionally large MHCs with 20 and 80 heme binding sites (**Figure 11, Figure 9—figure supplement 3**). Similarly large cytochromes in archaea are only found in *Geoglobus* and *Ferroglobus* (McGlynn et al., 2015). Both of these groups are capable of extracellular iron reduction (Slobodkina et al., 2009; Smith et al., 2015). The largest and best conserved of these MHCs were classified into three major groups (ANME MHC Type A, B, and C) based on sequence identity, conserved domains, and heme-binding motif distribution (**Figure 11**). The most notable feature of ANME MHC-A and B is the presence of predicted S-layer domains, which commonly make up the outer proteinaceous shell surrounding the cytoplasmic membrane in many archaea (Albers and Meyer, 2011). ANME MHC-B additionally contain an N-terminal domain free of heme-binding motifs and annotated as "peptidase M6-like domains". The presence of an S-layer domain in these cytochromes suggests their use for electron transfer through this outermost layer. In *M. acetivorans* the S-layer domain structure has been determined by x-ray crystallography and contains two subdomains

connected by a flexible linker (Arbing et al., 2012). In all ANME MHC with S-layer domains, this flexible linker also contains a single heme-binding motif, which would place a heme group within the plane of the S-layer. Both ANME MHC-A and B encode for C-terminal regions predicted to be transmembrane helices that may anchor them cytoplasmic membrane. We suggest that ANME MHC-A and B are functionally analogous to the MtrCAB complexes found in the outer membranes of the EET-capable gram-negative bacteria (**Figure 9A**) (Richardson et al., 2012). In ANME an alternative mechanism needs to be employed since the outermost layer is proteinaceous.

ANME MHC type C do not encode for an S-layer domain or C-terminal transmembrane helices, but do encode for a large N-terminal "pectin lyase-like domain". It is difficult to predict the function of these additional domains, however it is interesting to note that the pectin lyase fold typically occurs in proteins that attach to and/or degrade carbohydrates, and has been found in bacteriophage tail spike proteins for attachment to hosts (Latka et al., 2017). It is possible then that this domain is involved in recognizing the outer cell wall of partner bacteria. Some ANME genomes contained exceptionally large MHC that did not fall clearly into these three categories. ANME-3 HMMV2 for example contained a single polypeptide containing C-terminal features of ANME MHC-C and N-terminal features of ANME MHC-B (**Figure 11**). A fosmid assigned to ANME-2c ERB4 encoded the largest MHC in our dataset containing an S-layer domain, a pectin lyase-like domain, and 86 heme binding motifs.

In actively growing ANME-2c cultures, genes encoding these large ANME MHC were expressed at lower levels than the enzymes of the reverse methanogenesis pathway and some of the smaller MHCs leading the authors to conclude that these were of little importance to electron transfer in this culture (Krukenberg et al., 2018). If two proteins are expected to carry out the same function in a pathway,

transcription levels may be useful in determining which may be the dominant one operating under certain conditions. But this is not the case for the different classes of MHCs in the model of ANME EET presented here. The small MHCs are thought to serve as intermediates between the inner membrane and S-layer, as well as assist in conveying conductivity to the extracellular matrix, while the large S-layer proteins specifically provide a conduit through the S-layer. The small MHCs may cover longer distances, and would therefore be expected to be numerically dominant, but still could not make a functional EET system if no specific mechanisms existed for the short transfer through the S-layer to the extracellular environment. This is exactly the scenario observed in the electrogenic model bacterium *G. sulfurreducens* that requires MHC conduits through the outer membrane for EET formed by MtrCAB and other analogous porin-cytochrome systems (Richardson et al., 2012). While absolutely necessary for efficient EET, these genes show lower expression levels that those of some of the smaller soluble MHCs encoded in the genome (Jiménez Otero et al., 2018). We therefore find this previously reported transcript information to be consistent with the model of ANME-2 and 3 EET presented here in which the S-layer fusions play a key role (**Figure 9B**).

ANME-1 contained fewer MHCs and at most had 10 predicted heme-binding motifs, and these proteins did not contain any identifiable S-layer domains. From the available genomic data, it is unclear how these cytochromes fit into the cell structure of ANME-1. It was recently suggested that ANME-1 lack an S-layer due to the absence of predicted S-layer domain containing proteins (Timmers et al., 2017). However, the same analysis found no *Archaeoglobus* proteins with S-layer domains, yet the *Archaeoglobus* S-layer has been visualized and extensively characterized (Kessel et al., 1990). The S-layer domain in the large MHC from ANME-2 and 3 are only recognizable because they are homologous to the S-layer protein of *M. acetivorans* that was recently crystalized (Arbing et al., 2012). Before this study the S-layer

domain was only identified as a domain of unknown function (DUF1608). The ANME-1 ultrastructure is usually cylindrical and highly reminiscent of *Methanospirillum* which contain S-layers (Beveridge et al., 1991). How the MHC in ANME-1 fit into this cell structure will require further investigation. Gram positive bacteria are also capable of EET, and in some cases have been found to use more modestly sized MHC proteins (6-9 heme binding motifs) for electron transport through the outermost cell wall (Carlson et al., 2012). Such a model could very well apply to the cell wall and cytochromes encoded in the ANME-1 genomes.

Duplication of Cytochrome c *maturation machinery*

Cytochrome *c* proteins are matured on the outer face of the cytoplasmic membrane through a variety of systems, the most common in bacteria and archaea being the cytochrome *c* maturation (Ccm) system (Sanders et al., 2010). Ccm transports heme B to the positive side of energized membranes and mediates its attachment to the cytochrome *c* apoprotein (**Figure 12**). The Ccm system has been found in archaea closely related to ANME (Kletzin et al., 2015), and slight modifications to these systems in archaea have been investigated in detail (Allen et al., 2006). Most ANME genomes contained CcmABCEF and lacked CcmH a pattern observed previously in most other archaea including methanogens. One exception to this is the freshwater ANME-2d, which contains CcmH homologs similar to those found in *Ferroglobus* and *Geoglobus sp.* (Allen et al., 2006; Kletzin et al., 2015).

A very interesting feature of the ANME genomes reported here is the presence of multiple paralogs of CcmF. CcmF is the largest of the Ccm proteins, and is involved in the transfer and attachment of heme B to the cytochrome *c* apoprotein. One of these highly divergent CcmF paralogs that is conserved in ANME-2a, 2b, 2c, 2d, 3 and "*Ca.* Argoarchaeum" is found in every case next to the S-layer containing

ANME MHC-B (**Figure 11, 12B**). The diversification of CcmF paralogs in ANME and the clustering of one of the CcmF paralogs with ANME MHC-B suggests that special CcmF proteins may be needed to handle the extremely large apoproteins associated with these MHCs. The specialization of CcmF homologs to particular cytochromes *c* is not without precedent, a specific CcmF homolog NrfE is known to be required for insertion of heme B into NrfA at a modified heme binding site with a CxxCK motif (Lee et al., 2007). These additional CcmF homologs in ANME may be required due to the extreme size of the ANME MHC apoproteins or due to the presence of multiple modified heme biding motifs with increased amino acids between the cysteine residues (**Figure 11**).

Immediately upstream of the CcmF/ANME MHC-B genes in ANME-2a, 2b and 3 are two highly conserved genes also expected to be involved in MHC maturation (**Figure 12B, Figure 9—figure supplement 1**). One is a signal peptidase which is likely involved in the cleavage of the N-terminal signal sequence present on most of the encoded MHC proteins in ANME. The second is an integral membrane protein that has no automated annotations (labelled 8x TMH), and is found in every ANME group, the "*Ca.* Syntrophoarchaeum", "*Ca.* Argoarchaeum", and *Geoglobus/Ferroglobus*, but no other genomes based on BLAST searches of the NCBI nr database. While no specific motifs are clearly obvious, a large periplasmic loop is predicted that contains multiple highly conserved tryptophan residues which is a common feature of the heme handling proteins (HHP) family to which both CcmF and CcmE belong (Lee et al., 2007). This protein may have a role in heme handling in the maturation process of the MHC proteins in ANME as well. These two genes are also found in a cluster of cytochrome *c* genes highlighted in a previous metagenomic study of ANME-1 (Meyerdierks et al., 2010). Due to their similarity to peptidases and heme handling proteins we suspect these are not membrane anchors for the mature MHC proteins as previously posited. The unusual ANME-1 subclade consisting of GoMg4 and SA

contained no cytochrome *c* genes and also lacked CcmABCE, all CcmF paralogs, the signal peptidase, and the integral membrane protein described above. This provides an additional validation of the quality of our genome bins, as well as further highlighting the interesting biological differences in this subgroup of ANME-1. In conclusion, a key genomic difference between canonical methanogens and the ANME archaea are traits linked to direct interspecies electron transfer, especially large, diverse cytochromes and associated biosynthetic machinery. It should be noted that it is possible for engineered syntrophies to occur with methanogenic *Methanosarcinaceae* even in the absence of MHC, although the mechanistic details of how electrons are transferred in this system are lacking (Yee and Rotaru, 2020).

Anabolic pathways

In addition to energy metabolism reconstruction, the diversity of ANME genomes presented here allows for an evaluation of the anabolic pathways present in the different methanotrophic lineages. The assimilation of isotopically light (^{13}C-depleted) carbon into ANME lipids and bulk biomass was crucial to their initial discovery (Boetius et al., 2000; Hinrichs et al., 1999; Orphan et al., 2002, 2001b; Pancost et al., 2000), and is an important signal used for the interpretation of stable isotope studies of ancient carbonate systems. A better understanding of the precise biochemical pathways available for building this biomass will help better interpret these types of analyses, as well as the results of stable isotope probing experiments. Early experiments also highlighted how isotope signatures serve to identify the syntrophic cooperation between the consortia partners (Orphan et al., 2001b, 2001a). In this regard, long-term symbiotic interactions between organisms has been hypothesized to lead to lasting anabolic impressions on the organisms involved, specifically the adaptive loss of anabolic pathways such that syntrophic partners are left with complementary set of functions such as amino acid biosynthesis

(Embree et al., 2015; Morris et al., 2012). Below we highlight some of the key conserved features and differences between anabolic pathways present across the different groups of ANME archaea.

Anabolic C1 metabolism

C1 metabolism concerns the enzymes, reaction and cofactors mediating redox transformations of single carbon molecules and the ways these are incorporated into various biomolecules through the formation of carbon-carbon bonds. The single known anabolic C1 pathway present in all ANME is the Wood-Ljungdahl (reductive acetyl-Coenzyme A) pathway. The methyl branch of this pathway is the methanogenesis pathway described in previous sections. In this pathway, methyl groups attached to H_4MPT are combined with CO_2 and a pair of electrons from Fd^{2-} to make acetyl-CoA which can then be used in a wide variety of biosynthetic processes. This pathway ANME share with methanogens, some of which are autotrophic, i.e. derive their cell carbon from CO_2 reduction. Still, an important question for interpreting isotopic studies is whether there are other means of C1 assimilation and, if so, if these alternative mechanisms vary within and between ANME clades.

The H_4MPT cofactor of the Wood-Ljungdahl pathway was first discovered in methanogenic archaea and was originally thought to be an archaea-specific modified folate substituting in place of H_4F in all methanogens and their relatives for C1 carrying reactions. This view held until it was shown the H_4MPT was also present in methylotrophic bacteria (Chistoserdova, 1998). These bacteria contain both H_4MPT and H_4F and assessing roles of the C1 pools associated with each pathway for either catabolism or anabolism was a major challenge requiring decades of biochemical, genetic, and comparative genomic analysis to begin to fully understand (Chistoserdova et al., 2009). In methylotrophic bacteria H_4MPT appears to be the dominant catabolic oxidation pathway for formaldehyde generated by methanol

dehydrogenase, while H_4F can either serve as a catabolic oxidation pathway for certain methylated compounds or as a mechanism for transferring C1 moieties from formate into the anabolic serine cycle. Because of the nuanced understanding of C1 metabolism in methylotrophic bacteria and the apparent importance of maintaining separate C1 pathways in these organisms, we investigated the extent to which H_4F-bound C1 pools may be present in the ANME archaea (**Figure 13**).

In ANME-2a, 2b, 2d and 3, MetF is clustered with the genes for MetV, FolP, FolD, GlyA, and PurN, all of which are involved in H_4F-related anabolic C1 reactions (**Figure 13C**). This cluster of genes is also found in other closely related methanogens in the *Methanosarcinaceae* (Buchenau and Thauer, 2004). PurH1, encoded elsewhere in these genomes, also utilizes H_4F. A range of biochemical and isotope labelling studies found that these enzymes are involved in anabolic pathways and have demonstrated their specificity for H_4F in *M. barkeri*. Importantly, the C1 moieties in the H_4F pathway are not derived from free C1 compounds, as the critical enzyme formate H_4F ligase is not present in any ANME genome. Instead, in ANME the positions of compounds that are derived from methylene-H_4F will come from the C2 position of acetyl-CoA as was shown in *M. barkeri* (Buchenau and Thauer, 2004). Acetyl-CoA derived carbon passes through pyruvate to serine, and through the activity of serine hydroxymethyl transferase (GlyA), the carbon derived from the C2 of acetate is transferred from serine to H_4F to be used for biosynthesis of purines, thymine and possibly methionine (**Figure 13A**). The methyl group donor for methionine biosynthesis is likely a corrinoid protein, but its identity remains unknown in methanogens (Schroder and Thauer, 1999).

ANME-2c lacks a MetF of either the normal *Methanosarcinaceae* variety, or the ANME-1/uncultured archaeal version. Consistent with the lack of the *Methanosarcinaceae* variety of MetF, all ANME-2c

genomes additionally lack MetV, FolP, FolD, PurN and PurH1 (**Figure 13C**). This absence of H_4F-interacting genes is also true of all the ANME-1 genomes, save for MetV which is found alongside the ANME-1 MetF. This suggests H_4F is not used in ANME-1 or 2c for biosynthetic reactions (**Figure 13B**)

GlyA homologs are found in all ANME genomes. Since GlyA is expected to interact with H_4F in ANME-2a, 2b, 2d and 3, we conducted a deeper analysis of GlyA in order to find an explanation for their presence in ANME-2c and ANME-1 since they seemed to lack the other H_4F-related genes. Phylogenetic analysis of GlyA showed a clear differentiation between GlyA in ANME-1 and 2c vs. those in ANME-2a, 2b, 2d and 3 (**Figure 13D**). GlyA in methanogens without H_4F react with H_4MPT (Angelaccio et al., 2003; Hoyt et al., 1986), while the *Methanosarcinaceae* with H_4F contain GlyA active towards H_4F (Buchenau and Thauer, 2004). ANME-1 and 2c GlyA sequences cluster with the GlyA that have been biochemically characterized to interact with H_4MPT, while GlyA from other ANME cluster with those of *Methanosarcinaceae* and would be expected to react with H_4F (Lin and Sparling, 1998). This pattern of presence/absence of genes and phylogenetic affiliation of homologs suggests that H_4F is present in the ANME-2a 2b 2d and 3, H_4F is present and used in the same manner as *M. barkeri*. In ANME-1 and 2c we expect that GlyA is necessary to produce glycine from serine, but that the resulting C1 moiety is shuttled back into the H_4MPT-bound C1 pool.

PurH1 and PurN which are absent in ANME-1 and ANME-2c utilize formyl-H_4F as a source of C1 moieties for two of the carbon atoms in the biosynthesis of purines (**Figure 13A**). An interesting question is how ANME-1 and 2c are able to grow without these enzymes. Alternate enzymes, PurT and PurP/PurH2, can carry out these steps, but use free formate as the C1 donor (Brown et al., 2011).

Notably, ANME-2c contained both PurT and PurH2 which are absent in the ANME-2a, 2b, 2d and 3, while ANME-1 genomes contained PurH2 but lacked PurT. A lack of both PurT and PurN has been observed in *Archaeoglobales* and *Methanobacteriales*, which suggests that there is an as yet unidentified enzyme catalyzing this step since both are known to synthesize their own purines (Brown et al., 2011). This distribution of genes involved with purine biosynthesis indicates at least an anabolic role for formate in ANME-1 and 2c metabolism.

Another important anabolic C1 reaction is carried out by thymidylate synthase, a crucial step in the synthesis of the thymidine base of DNA, which catalyzes the methylation of deoxyuridine monophosphate (dUMP) to produce deoxythymidine monophosphate (dTMP). This C1 moiety is often derived from 5,10-methylene-H_4F, and this reaction can be carried out by two non-homologous thymidylate synthase proteins, known as ThyA and ThyX (Myllykallio, 2002). The reactions catalyzed by these enzymes are slightly different, with ThyA using the H_4F itself as an electron donor to reduce the methylene to methyl, producing a dihydrofolate product. ThyX in contrast utilizes NAPDH in the reaction, leaving H_4F in the tetrahydro oxidation state.

The genomes of ANME-2a, 2b, 2c, 2d, and 3 contain ThyA, whereas most ANME-1 have ThyX. The only exception to this is the small ANME-1 subgroup comprised of genomes GoMg4 and SA that contain ThyA instead of ThyX, adding to the list of unique features of this ANME-1 subclade. Phylogenetic reconstruction of ThyA homologs revealed a similar split between ANME with and without H_4F; ANME-2a, 2b, and 3 cluster with their close methanogenic relatives, while ANME-1 and 2c cluster with *Methanocellales*. Curiously ANME-2d also cluster with ANME-1 and 2c although they contain all the other H_4F interacting proteins. Much less is known about ThyA biochemistry in

methanogens. Labeling and biochemical studies have not been carried out to the same degree as with GlyA, making it is difficult to predict the cofactor specificity for either of these ThyA clusters. Due to the presumed lack of H_4F in ANME-1 and 2c the ANME-1, 2c, and 2d cluster likely utilizes H_4MPT, while the ANME-2a, 2b, and 3 cluster utilizes H_4F (**Figure 13A, B**).

The difference between using ThyX or ThyA could be particularly useful because thymidylate synthase has recently been the subject of great interest as a druggable target in *Mycobacterium tuberculosis*. ThyX is thought to be essential for *M. tuberculosis* while humans contain ThyA, therefore various ThyX inhibitors are being investigated for clinical antimicrobial uses (Basta et al., 2012). These ThyX inhibitors could be valuable tools for ANME ecology research if they are shown to be specific inhibitors of ANME-1 in order to determine the contribution of ANME-1 to methane oxidation in mixed ANME communities.

Apparent amino acid prototrophy

Obligate coupling between the energy metabolisms of syntrophic organisms has been suggested to lead to additional metabolic dependencies such as amino acid or vitamin auxotrophies through adaptive gene loss (Embree et al., 2015; Morris et al., 2012). Yet, the genomes of all ANME clades contain a near complete complement of genes predicted for synthesizing all amino acids, including many of the pathway modifications described from cultured methanogens (**Figure 13—figure supplement 1**). Of the enzymes that are widespread in related methanogens, only phosphoribosyl-ATP diphosphatase (EC 3.6.1.31), a step of histidine biosynthesis is not found in any other ANME genomes, with the exception of ANME-3 (**Figure 13—figure supplement 1**). The gene responsible for this step is also absent in *Archaeoglobus fulgidus* and *Nitrosopumilus maritimus*, both of which are capable *de novo* histidine synthesis,

suggesting other unknown mechanisms exist for completing this step of histidine biosynthesis (Fondi et al., 2009).

Other isolated steps in amino acid synthesis were also unannotated in ANME genomes, however they were also missing in related methanogens known to complete these steps in an as yet uncharacterized manner. ANME genomes appear to lack phosphoserine transaminase but this is possibly complemented by a broad specificity aspartate aminotransferase as has been described in other methanogens (Helgadóttir et al., 2007). Aromatic amino acids appear to be produced via the 6-deoxy-5-ketofructose-1-phosphate (DKFP) pathway to produce shikimate (Porat et al., 2006; White, 2004; White and Xu, 2006), however DKFP synthase was not detected in most of the ANME genomes or in many of the *Methanosarcinales*, again suggesting an as yet unidentified alternate gene may be used in this process (**Figure 13—figure supplement 1**). Methionine appears to be synthesized using the partially described pathway converting aspartate-semialdehyde directly to homocysteine, though not all enzymes are known for this pathway (Allen et al., 2015). The final steps of tyrosine and phenylalanine synthesis did not have annotated genes in many methanogens or ANME genomes but may be complemented by a broad specificity amino acid transaminase. Finally, histidinol-phosphatase (EC3.1.3.15) that is also part of histidine biosynthesis is unannotated in all of the ANME and methanogen genomes, suggesting an alternate, undiscovered gene operating at this step.

Incomplete partial TCA cycles for 2-oxoglutarate synthesis

Many biomolecules including porphyrins and amino acids are synthesized from 2-oxoglutarate, and this intermediate comes from the TCA cycle beginning with acetyl-CoA sourced from CH_3-H_4MPT and CO_2 via the acetyl-CoA decarbonylase/synthase complex (ACDS). In methanogens, a complete TCA cycle

is not needed. Instead, one of two partial pathways are present to produce 2-oxoglutarate for biosynthesis: either the partial TCA cycle operates in the oxidative direction passing through isocitrate, as found in the *Methanosarcinales,* or it runs in the reductive direction through succinate, as in *Methanococcus* (Simpson and Whitman, 1993)(**Figure 14**). Surprisingly, ANME-2a and 2b largely contain the enzymes of the reductive partial TCA cycle, diverging from their close *Methanosarcinaceae* relatives (**Figure 14**). In contrast, the ANME-1 and 2c universally lack some of the genes in the reductive pathway, but contained the three genes catalyzing the reactions in the partial oxidative pathway. This pathway would involve ATP citrate lyase operating in the direction of ATP production, which is not typical. While some ANME-2a come close to having a complete TCA cycle, no single genome was found that contains all steps in the pathway, supporting the notion that these enzymes are more used as a means to make an important biosynthetic intermediate, not for running a complete reverse TCA cycle for carbon fixation. The ANME-3 genomes lack the complete set of genes for either pathway, and it is currently unclear how they can produce 2-oxoglutarate.

Additional ANME features of interest

In addition to the catabolic and anabolic pathways described above, we identified a number of other genomic features that appear interesting based either on their unique presence in ANME as compared to closely related methanogens, or their potential for being important in other aspects of the ANME-SRB syntrophy.

Nitrogenase in ANME

A phylogeny of the large subunit of nitrogenase, NifD, indicates that the ANME nitrogenases of the "methane seep group" form a monophyletic clade (**Figure 15**). The closest relatives of this group are from *Roseiflexus castenholzii*, which was also reported for *nifH* gene clones (Dekas et al., 2009). This cluster is notably divergent from the NifD in methanogenic *Methanosarcinales* capable of N_2 fixation (**Figure 15**).

The ANME NifD, as well as *R. castenholzii* and *Endomicrobium proavitum*, all contain the conserved cysteine and histidine residues for binding the P-cluster and FeMoCo ligands (McGlynn et al., 2012), which suggests they may be functional in nitrogen fixation, consistent with *nifH* transcript expression in the environment and correlated FISH-NanoSIMS analysis of $^{15}N_2$ assimilation (Dekas et al., 2009). All gene clusters of the methane seep group exhibit the gene order $NifHI_1I_2DK$, and lack the *nifN* and *nifE* genes that function as molecular scaffolds for the maturation of the FeMo cofactor (**Figure 15—figure supplement 1)**. These genes were traditionally thought to be crucial for a functioning nitrogenase (Dos Santos et al., 2012; Hu and Ribbe, 2016). However, a recent report of nitrogen fixation in *E. proavitum*, which also lack *nifEN*, suggests that these genes are not strictly necessary for nitrogen fixation (Zheng et al., 2016).

It was unexpected to find nitrogenases of the "methane seep group" in many of our ANME-1 genomes. Clone libraries of *nifH* amplicons from methane seeps were used to define four major groups present and expressed in these habitats referred to as: methane seep group, methanosarcina-like group, group III and group IV (Dekas et al., 2016). The methane seep group is composed of sequences almost exclusively found in seep environments and was assigned to members of the ANME-2 based on multiple lines of evidence (Miyazaki et al., 2009; Pernthaler et al., 2008). Group IV were detected in ANME-1 fosmids and shown to be poorly expressed by metatranscriptome analysis. This low expression coupled with the fact that group IV NifH were proposed to be involved in F_{430} biosynthesis instead of nitrogen fixation (Staples et al., 2007) led to conclusion that ANME-1 are likely incapable of nitrogen fixation (Meyerdierks et al., 2010). Group III NifH sequences are largely composed of sequences similar to those found in *Deltaproteobacteria*, likely corresponding to the ANME's syntrophic partners (Dekas et al., 2016). The *Methanosarcina*-like group is very similar to sequences found in genomes of the *Methanosarcina* genus,

and they were expected to come from low abundance methanogens occasionally present in methane seep environments. The wide distribution of methane seep group nitrogenase in addition to nitrogenases belonging to the group IV in ANME-1 suggest that these organisms may also be capable of nitrogen fixation.

The monophyletic nature of the nitrogenase genes in ANME is notable in the context of the paraphyletic nature of ANME organisms. This form of nitrogenase may have been present in the common ancestor of all ANME, and was subsequently lost in all methanogenic members of the *Methanosarcinaceae* and then replaced with a phylogenetically distinct nitrogenase in some lineages. Alternatively, separate horizontal gene transfer events could have taken place to insert this nitrogenase into the various ANME clades which would seem to imply a strong selection for this specific nitrogenase clade in the ANME metabolism. Notably, ANME-3 is the only ANME clade lacking a nitrogenase suggesting a requirement for fixed nitrogen, such as porewater ammonium (Bowles and Joye, 2011).

ANME-1 genomes harbor many FrhB/FdhB/FpoF paralogs

Since F_{420} is an electron carrier expected to be used by ANME we examined the diversity of genes predicted to be involved in F_{420} redox reactions. A number of proteins that interact with F_{420} have been discussed above, and many of them share homologous subunits that carry out F_{420} redox chemistry such as FrhB (F_{420} dependent hydrogenase), FdhB (F_{420} dependent formate dehydrogenase) and, FpoF/FqoF, etc. Other enzymes utilizing this FrhB-like domain are known such as F_{420}-dependent sulfite reductase (Fsr), which has recently been examined in detail in ANME (Yu et al., 2018). Surprisingly, our examination of this protein family revealed many additional unknown paralogs, particularly in ANME-1 (**Figure 16**). These paralogs were generally monophyletic within the ANME-1, suggesting duplication

and neofunctionalization within the order *Syntrophoarchaeales* (**Figure 16—figure supplement 1)**. The gene clusters containing these FrhB paralogs were often well conserved, and seemed to contain genes that code for enzymes not generally expected to interact directly with F_{420}. For example, FrhB6 is mainly found in gene clusters containing pyruvate ferredoxin oxidoreductase, the enzyme responsible for the reductive carboxylation of acetyl-CoA to form pyruvate. Similarly, FrhB7 was found in gene clusters with CODH alpha and epsilon subunits. These subunits of the larger CODH/ACS complex are responsible for the carbon monoxide dehydrogenase activity (Gong et al., 2008).

If these two enzyme systems form complexes with the novel FrhB paralogs as this gene clustering suggests, then they may be operating only in the oxidative direction, using electrons from the CO/CO_2 couple or the oxidative decarboxylation of pyruvate to reduce $F_{420}H_2$. Biochemical precedent seems to indicate that $F_{420}H_2$ cannot be an electron donor for these reactions under physiological conditions. It was initially thought that pyruvate oxidoreductase responsible for the anabolic formation of pyruvate could utilize $F_{420}H_2$ as its physiological electron donor based on experiments with crude cell extracts (Zeikus et al., 1977). Subsequent studies that purified the pyruvate oxidoreductase enzyme found F_{420} to not be a substrate as an electron donor or acceptor (Bock et al., 1996). Instead, pyruvate oxidoreductase interacted with ferredoxin, a finding consistent with previous results from bacteria. This result also is also more consistent with bioenergetic arguments, the pyruvate/acetyl-CoA+CO_2 redox couple is expected to be around -500mV and therefore should not be able to be reduced by $F_{420}H_2$ with a midpoint in the vicinity of -340mV. Thus the pyruvate dependent reduction in crude cell extracts was attributed to secondary reduction of F_{420} by Fd (Bock et al., 1996). The gene clusters here may represent a dedicated system of F_{420} reduction associated with the breakdown of multi-carbon compounds. Whether this

system is used in the breakdown of carbon storage molecules, or is an important part of ANME catabolism, remains to be determined.

Additional noteworthy ANME-specific FrhB paralogs include FrhB10, which are found in gene clusters with various subunits of glutamate synthase, potentially representing a third putative group of F_{420}-dependent glutamate synthases (Susanti and Mukhopadhyay, 2012). FrhB8 are found in gene clusters with divergent NuoF homologs and thioredoxin genes, and are almost exclusively found in ANME groups (**Figure 16—figure supplement 1**). While our limited transcriptome data shows many of these FrhB homologs to be expressed at moderate levels, FrhB7 in ANME-1 GB60 is an exception, with expression levels near those of the methanogenesis pathway (**Supplementary file 2**). As with the diversity of HdrA-containing gene clusters, this expansion of FrhB-containing gene clusters in the ANME-1 points to a wide range of novel electron transport capabilities in ANME-1.

Extensive Dockerin/Cohesin domain-containing proteins

Recently a protein domain study of three marine ANME genomes reported that ANME-1 and ANME-2a contained a surprising number of proteins predicted to have dockerin or cohesin domains (Timmers et al., 2017). These domains are best known from their role in the construction of the cellulosomes which are large, multiprotein complexes that bind and degrade extracellular cellulose in *Clostridia* and other cellulose degrading bacteria (Bayer et al., 2004). Dockerin and cohesin domains form strong bonds with one another, and by encoding multiple sets of complementary dockerin and cohesin-containing proteins many copies of enzymes with cellulolytic activity can be linked to other proteins containing cellulose binding domain as well as anchored to the cell surface. These multiprotein complexes can be

bound to the cell by anchor proteins that contain cohesion and S-layer homology (SLH) domains or via covalent linkage to lipids through the action of sortases.

Cohesin and dockerin domains have been found in diverse bacteria and archaea that are not thought to be involved with cellulose degradation, suggesting functions beyond the well characterized ones in clostridia (Peer et al., 2009). Early work found proteins containing dockerin and cohesin domain in *Archaeoglobus* and these archaeal versions have been verified to perform specific strong dockerin-cohesin bonds (Bayer et al., 1999; Haimovitz et al., 2008). However archaea that contained dockerin or cohein domains did not contain large so-called "scaffoldin" proteins with multiple copies of cohesin domains which are required for the formation of large multimeric complexes (Peer et al., 2009). Since dockerin and cohesin domains simply facilitate the linkage of functional domains, the implications of these domains are not clear simply by their presence in the genome.

Dockerin and cohesin domain-containing proteins in bacteria vary considerably between closely related species in the number and identity of additional domains, and the proteins in ANME are similarly variable, but a few consistent trends are apparent. In members of ANME-2a, ANME-2c and ANME-1 there are proteins that consist of a dockerin domain and multiple cohesin domains, up to as many as five. These scaffoldin-like proteins therefore have the ability to localize multiple copies of whatever proteins contain their cohesins' complementary dockerin domains. Common dockerin containing protein in ANME are the periplasmic substrate-binding components of ABC transporters of various types, such as the nickel/dipeptide/oligopeptide (NikA/OppA) system and the TroA system used for metal ion uptake (Berntsson et al., 2009; Zheng et al., 2011). In most ANME-2c genomes there are proteins encoding S-layer domains and dockerin domains, suggesting dockerin/cohesin pairing may be

important for attaching functional proteins to the outer layer of the cell and possibly mediating interactions with bacteria. Using experimental methods to determine which dockerin and cohesin domains bind to each other will lead to a better understanding of composition of the ANME extracellular space. A complete list of ANME proteins containing dockerin or cohesin domains can be found in **Supplementary file 3**.

Phage-like protein translocation structures

Phage-like protein translocation structures (PLTSs) are large multiprotein complexes that share structural and functional similarities to type VI secretion systems, pyocins, and the *Serratia entomophila* antifeeding prophage described in a broad survey of microbial genomes based on sequence identity and synteny (Sarris et al., 2014). This broad family of related complexes have a wide variety of functions, from neighboring cell lysis, to morphologic transformation of targeted eukaryotic cells (Shikuma et al., 2014), stabilizing symbiotic interactions (Speare et al., 2018) and mediating sibling conflict in the multicellular aggregate bacterium *Myxococcus xanthus* (Troselj et al., 2018). PLTSs are anchored in the cytoplasmic membrane and upon contraction of the sheath proteins extrude a protein complex with a sharp spike that penetrates neighboring cell membranes and can deliver various effector proteins (Shneider et al., 2013). Recent reviews of the known structure and function of these complexes highlight the conserved features which include phage baseplate-like proteins, sheath proteins, AAA+ ATPases, LysM-motif proteins for peptidoglycan binding, and VgrG and PAAR-domain spike proteins (Taylor et al., 2018).

Gene clusters related to these systems were identified in all ANME groups with gene synteny similar to that of *Methanomethylovorans hollandica* that was described in a recent review (Sarris et al., 2014) (**Figure**

17). This gene cluster was relatively well expressed in ANME-2c (**Supplementary file 2**), and could potentially play an important role in the symbiosis between ANME and partner bacteria, or alternatively defend AOM consortia from invasion. Interestingly, PLTS clusters from closely related ANME-2a and 2b were very different from one another, with the ANME-2b sequences much more closely related to the ones found in ANME-1 (**Figure 17C**). In the two ANME-2d genomes available, one encoded the version like ANME-2a, ANME-2c, and ANME-3, while the other ANME-2d genome encoded a version that was very similar to ANME-1 and ANME-2b. This difference is most apparent in the PAAR domain spike proteins, with ANME-2b much more similar to the PAAR domain proteins encoded in ANME-1 (**Figure 17C**). It is interesting that such closely related ANME groups encode such different PLTS systems, and suggests that ANME-1 and 2b may have similar interactions with organisms (possibly syntrophic SRB partners) that are quite different than those in association with ANME-2a, 2c and 3.

Discussion

The evolution and conserved metabolic features of marine ANME archaea

It remains a microbiological enigma that methanogens, i.e. those that produce methane as a waste product during their catabolic metabolism, and methanotrophs, i.e. those which gain energy from the consumption of methane, use the same set of core genes, but can not be reversed Our investigation of this expanded set of ANME genomes revealed key features present in ANME that are absent in their methanogenic relatives. A unifying feature of the ANME archaea appears to be a multitude of multiheme *c*-type cytochromes which are rare in methanogens. This expansive repertoire of MHCs, some harboring 30+ heme binding domains in ANME-2 and 3, specific cytochrome *c* maturation machinery, and associated conserved hypothetical proteins are features only shared by certain members of the *Archaeoglobales* that are known to conduct EET like *Ferroglobus*. The large ANME-MHCs with S-layer

fusions and the repurposing of the cytochrome *b* subunit of membrane bound hydrogenases into a potential methanophenazine:cytochrome *c* oxidoreductase complex (Mco) remain some of the clearest examples of bioenergetic novelty in ANME compared to their methanogenic relatives with which they share the core methanogenic pathway. The study of EET in symbiotic associations is relatively new, and there are no specific marker proteins yet identified that definitively confer EET ability. There is however a general pattern in that the genomes of EET-capable organisms contain many, often exceptionally large, MHCs, as well as systems enabling the transfer of electrons through the various outer layers of the cell. ANME appear to contain all of these genomic features, and this clear EET potential represents an important physiological difference between ANME and methanogens.

Another common feature of ANME genomes that set them apart from their methanogenic counterparts is the high abundance and diversity of soluble Hdr complexes with distant formate dehydrogenase (FdhAB) homologs. This is especially true for ANME-1 that lack certain features often found within *Methanosarcinaceae* including HdrDE and the Rnf complex in most cases. It is likely that these soluble Hdr complexes are involved in electron recycling and transfer processes in at least ANME-1, but also ANME-2 and 3 based on their conservation and expression levels. So far, experimental evidence suggests that formate is not an electron carrying intermediate in AOM, and the canonical formate transport protein (FdhC) is absent. Hence, the specific substrate and function of these well-conserved complexes remains a fundamental question in ANME metabolism remaining to be addressed. Regardless of the nature of the electron acceptor utilized by these Hdr complexes, the reverse methanogenesis model requires that their combined effect is a net oxidation of CoM-SH/CoB-SH, presumably through electron confurcation.

This expanded set of ANME genomes also helps us better understand the evolutionary history of the ANME. The paraphyletic nature of ANME suggests that the transition between methanogenic and methanotrophic metabolisms has occurred multiple times. The order of these events is not immediately evident, as multiple scenarios could lead to the observed pattern of metabolism and gene content. This is particularly true in the case of ANME-3, which contains multiple features identified in ANME-2 genomes including large MHC proteins, specific duplicated CcmF proteins, the *b*-type cytochrome associated with Rnf, divergent FdhAB associated with HdrA genes, among others. The presence of these bioenergetic genes within ANME-3, which appear to be the most recently diverged from within the *Methanosarcinaceae*, strongly suggests that this repertoire is critical to the methane oxidation phenotype. With the present data we do not believe there is a clear parsimony-based argument to distinguish between the possibilities that these ANME-specific features represent a gain of function through horizontal gene transfer into the ancestor of ANME-3, or whether they were lost in the methanogenic members of the *Methanosarcinaceae*.

It has long been assumed that ANME evolved from methanogenic ancestors and our current phylogenies support this scenario, as many groups of hydrogenotrophic methanogens emerge from the archaeal tree before the rise of ANME-1 which is the deepest branching ANME group (**Figure 1**). With this expanded set of ANME genomes, another interesting evolutionary possibility presents itself: the entire group of methylotrophic methanogens within the *Methanosarcinaceae* may be derived from a methanotrophic ancestor. These methanogens run six of seven steps of the methanogenesis pathway in reverse. This fact was used early on to suggest "reverse methanogenesis" was reasonable for ANME, since they would only have to reverse one additional step. The order of our discovery of these metabolisms may not reflect the order with which these metabolisms evolved, and the complete reversal

of the methanogenesis pathway may date back to the last common ancestor of ANME-1 and the other ANME. In this evolutionary scenario, methylotrophic methanogenesis in the *Methanosarcinaceae* evolved by the simple acquisition of methyl transferases and the subsequent loss of ANME-specific systems of EET. Instead of transferring electrons from methane oxidation to the outside of the cell, the *Methanosarcinaceae* electrons are funneled from methyl group oxidation back into the cytoplasm to reduce heterodisulfide.

The "Methanoalium" group of ANME-1 and the potential for methanogenesis in ANME

The clade within ANME-1 represented by the SA and GoM4 genomes is an exceptional group which requires further detailed study. This group has previously been referred to as the "freshwater" ANME-1 clade and has been found in 16S rRNA gene surveys of various terrestrial and marine environments (Brazelton et al., 2010, 2006; Takeuchi et al., 2011). A recent report constructed the first genome from this group of ANME-1 from a Tibetan hot spring (THS) (Borrel et al., 2019). We recognized similar characteristics between the THS genome and the GoM4 and SA genomes reconstructed here, including the bacterial-type Rnf and the absence of MHC. Although not mentioned previously, we find the same MvhA-type hydrogenase in the ANME-1-THS genome and an absence of cytochrome maturation machinery. As a final check on the novel characteristics of this unusual ANME-1 clade, we reconstructed one additional genome from a recent metagenomic dataset of the Lost City hydrothermal field. This site has reported 16S rRNA genes and McrA sequences belonging to this freshwater ANME-1 clade (Borrel et al., 2019; Brazelton et al., 2010, 2006; Kelley et al., 2005). This additional Lost City ANME-1 MAG was 88% complete with 2% contamination and notably had all the hallmark genomic features observed in GoM4, SA, and THS (**Figure 18**).

A number of investigations of ANME-dominated environments have concluded that some ANME lineages, particularly ANME-1, maybe have net methanogenic capabilities (Bertram et al., 2013; Beulig et al., 2019; House et al., 2009; Lloyd et al., 2011). Genomic analysis alone cannot falsify the hypothesis that ANME archaea can gain energy from methanogenesis, due to the similarities in the main methanogenic pathway enzymes. Yet incubations of methane seep samples with methanogenic substrates has only succeeded in stimulating the growth of low abundance traditional methanogens, never ANME (Wegener et al., 2016). The evolution between methanotrophy (e.g. ANME-2a, 2b, and 2c), and methanogenesis (canonical members of the *Methanosarcinaceae*) as well as the apparently recent transition to back to methanotrophy in ANME-3 appears to include biochemical innovation and genomic adaptation which occurs on evolutionary timescales, but not ad hoc in the environment. Likely, in a group as old and diverse as the ANME-1 archaeal order (representing at least 6 genera in our genomic collection), such a transition may have also occurred. We hypothesize that this clade of ANME-1 (SA, GoM4, THS, and LC) that encode for hydrogenases and lack *c*-type cytochromes may be bona fide methanogens. However, the little evidence currently available for representatives of this clade from terrestrial subsurface suggests that this group may carry out methane oxidation as well (Takeuchi et al., 2011). Whether this group occurs in association with a syntrophic partner is currently unknown. What is clear is that this clade's genomic content and environmental distribution sets them well apart from other ANME. We propose the genus "*Ca.* Methanoalium" for the "freshwater ANME-1 clade" in recognition of the differences in their metabolic potential and unusual environmental distribution (see **Supplementary file 1** for etymology).

Anabolic independence of the ANME archaea from their syntrophic partner

The tight metabolic coupling between syntrophic ANME and SRB could be expected to result in reductive biosynthetic pathway loss in one of the partners as predicted in the "Black Queen Hypothesis" (Morris et al., 2012). Unfortunately, the partner SRB have not yet been cultured, and thus it remains unknown if these show specific genomic adaptations to a consortia lifestyle. However, we do not observe obvious evidence of this occurring at the level of major ANME lineages for amino acids. If this sort of reductive evolution does occur in the ANME-SRB symbiosis, it may be occurring at the species or strain level, and would be very difficult to confidently detect in coarse-grained genomic analysis with partial genomes. More detailed investigations of the question of complementary biosynthetic pathways between syntrophic partners will be best studied in sediments or enrichment cultures that are dominated by a single ANME-SRB pairing, where near complete genomes of both partners can be generated. But, on the broad scale of ANME evolution to their syntrophic lifestyle, there does not seem to be a concerted loss of anabolic independence in any of the major lineages. This result is distinct from that described in some other syntrophic communities such as those performing the anaerobic oxidation of hexadecane (Embree et al., 2015).

Biogeochemical and microbiological consideration of ANME carbon signatures

Bacterial methylotrophs and methanotrophs have traditionally been defined as organisms that derive both their carbon and energy from the oxidation of C1 compounds or methane, respectively (Hanson and Hanson, 1996; Lidstrom, 2006). ANME were called methanotrophs because it appeared that their energy metabolism was based on methane oxidation, and the ^{13}C-depleted isotope signature of their lipids were thought to be evidence of the assimilation of ^{13}C-depleted methane carbon. Biogeochemical studies further show that the respiration of methane to CO_2 dominates carbon turnover in most ANME

environments, and thus that the dissolved inorganic pool of carbon is mostly methane-derived. As a consequence, the carbon assimilated for biomass production is only a few percent of the methane oxidized and thus also methane-derived from the biogeochemical perspective (Nauhaus et al., 2007; Treude et al., 2007). However, in vivo isotope-probing experiments showed that ANME biomass is effectively labelled by $^{14}CO_2$ or $^{13}CO_2$, leading to the conclusion that ANME assimilate a mixture of CO_2 and CH_4 (Treude et al., 2007; Wegener et al., 2008), or almost only CO_2 (Kellermann et al., 2012). Short pulse-chase experiments with $^{14}CO_2$ and $^{14}CH_4$ also support this latter interpretation (Wegener et al., 2016).

Microbiological considerations for the classification of this C1 metabolism were best articulated by Leadbetter and Foster in the context of aerobic methane oxidizing bacteria (Leadbetter and Foster, 1958):

> *"Although there is no universally accepted definition of the nature of the autotrophic mode of life, the ability to grow at the expense of CO_2 as the exclusive source of carbon for cell synthesis remains as the cardinal consideration in the concept of an autotroph.*
>
> *The question needing study is whether, during growth on methane, these bacteria dehydrogenate the methane molecule to CO_2 and "active" hydrogen, following this with a reductive assimilation of the CO_2 with "active" hydrogen."*

The dominant bacterial methanotrophs in most environments assimilate methane-derived carbon at the oxidation state of formaldehyde for a large portion of their biomass using either the serine or ribulose monophosphate pathway (RuMP) (Lidstrom, 2006). Some organisms have been discovered that use the

oxidation of C1 compounds for their source of energy, but use a Calvin-Benson-Bassham (CBB) cycle for the fixation of CO_2. These organisms presented a problem for nutritional nomenclature because the definition of methylotroph/methanotroph normally involves the incorporation of cell carbon. For methylotrophs, Paracoccus denitrificans uses a CBB cycle when growing on C1 compounds, and these historically have been referred to as an "autotrophic methylotroph" (Anthony, 1982; Lidstrom, 2006). More recently, the methane-oxidizing bacteria *Methylacidiphilum fumariolicum* and "*Candidatus* Methylomirabilis oxyfera" have been found to operate a CBB cycle for carbon assimilation from CO_2 (Khadem et al., 2011; Rasigraf et al., 2014). These authors have taken a similar naming scheme, referring to *M. fumariolicum* as an "autotrophic methanotroph" (Khadem et al., 2011). These examples from bacteria can be transferred to the ANME, suggesting that the methanotroph label should be conserved, and the only question remaining is whether to adopt the "autotroph" label as well.

The predominant labelling of ANME biomass by $^{14}CO_2$ or $^{13}CO_2$ is intriguing, and suggest that they could be autotrophs or mixotrophs as some methanogens (Berg et al., 2010; Weimer and Zeikus, 1978). So far, all known ANME genomes encode solely the Wood-Ljungdahl pathway for carbon fixation that is also used for methane oxidation. One option to explain these isotope labelling results is that ANME are autotrophic methanotrophs, and have an additional, so far undetected CO_2 fixation pathway, so that anabolic and catabolic C1 reactions can run in opposite directions simultaneously, such as the CBB cycle in *M. fumariolicum*. Yet we have not found evidence for a second carbon fixation pathway in the ANME. Another possibility is that the C1 enzymes, C1 pools and cofactors used in the catabolic and anabolic pathways are separated spatially within the ANME cells. This could potentially be achieved by having two complete sets of paralogs of the enzymes of the Wood-Ljungdahl pathway that could be used in different directions and trafficking these paralogs to different compartments within the cell. While some

multiple paralogs of the methyl branch of the Wood-Ljungdahl pathway are spread sporadically throughout the ANME, we did not find any conserved complete second paralogous copy of all the enzymes of the archaeal Wood-Ljungdahl pathway. Another possibility could be a complete assimilatory Wood-Ljungdahl pathway utilizing H_4F as found in bacteria, but the H_4F-interacting enzymes found in some ANME-2 and -3 would not allow them to activate CO_2 via formate, and are instead are expected to be involved in specific anabolic reactions well downstream of acetyl-CoA production. If the ANME archaea are autotrophs, it remains to be determined how this is carried out from a biochemical perspective.

Another plausible explanation for the predominant labelling of ANME biomass by $^{13}CO2$ could be intense carbon back flux on the C1 branch on the Wood-Ljungdahl pathway to the point of CH_3-H_4MPT, against the net flux of carbon. Metabolisms operating at close to equilibrium can present challenges to isotope labeling studies because the forward and reverse fluxes through that pathway approach equal values (Holler et al., 2011). AOM operates as close to equilibrium as is thought to be possible to support life, and most of the steps of the methyl branch of the Wood-Ljungdahl pathway are essentially at equilibrium (Thauer, 2011). Under these conditions it seems quite possible that the isotope label of the DIC in these incubations could back flux up through the methyl branch of the Wood-Ljungdahl pathway to the point of CH_3-H_4MPT, even if the net reaction is in the oxidative direction. Further experiments are required to distinguish between these possible explanations, and determine whether ANME should be referred to as "autotrophic methanotrophs" or simply "methanotrophs".

MetF, F_{420}-dependent NADP reductase and electron bifurcation complexes

There remains no clear explanation for what evolutionary or bioenergetic pressures may have led to the loss of Mer and presumed replacement by MetF in ANME-1. An important challenge posed by this switch is that MetF interacts with NADPH instead of $F_{420}H_2$, and it is not immediately apparent how electrons from NADPH are harnessed for energy production in ANME-1. In one sense, this is similar to the case of methanogens that use multicarbon alcohols as electron donors for CO_2 reduction (Berk and Thauer, 1997). Some of these methanogens contain alcohol dehydrogenases that utilize F_{420} as the electron acceptor, and therefore the electrons liberated from the alcohols as $F_{420}H_2$ can immediately be used in the methanogenesis pathway. Other methanogens however contain alcohol dehydrogenases that are NADPH specific and so a secondary oxidoreductase is needed to transfer NADPH electrons onto F_{420}. It was discovered that in this latter group of methanogens, high activities of $F_{420}H_2$:NADP oxidoreductase (Fno) could be found, and the protein responsible was purified, characterized and sequenced (Berk and Thauer, 1998, 1997). If an Fno enzyme was expressed at sufficiently high levels in ANME-1 then the electrons on NADPH derived from MetF could be fed back into the F_{420} pool for catabolism, just as in the alcohol-oxidizing methanogens.

Remarkably, ANME-1 universally lack homologs of Fno although they are present in all other ANME and methanogens we examined. In most methanogens Fno is present for the reverse reaction, the production of NADPH for anabolic reactions. This conspicuous absence in ANME-1 of what seems to be a nearly universally conserved mechanism of hydride transfer between the $F_{420}H_2$ and NADPH pools is noteworthy on its own, but is particularly interesting when considered in the context of the Mer/MetF switch. If ANME-1 produce NADPH from MetF, they will have more electrons than they need for anabolism through this reaction since this amounts to a quarter of all methane-derived electrons. If this

is how MetF is being utilized in ANME-1, then how these NADPH electrons are transferred back into an energy conservation pathway is an important open question. Most ANME-1 also contain a homolog of what we have designated HdrA9, 10 or 11, which constitute HdrA with large insertions with high sequence identity with NfnB. This subunit interacts with NADPH in the well characterized Nfn systems (Demmer et al., 2015), so if MetF produces NADPH, these modified HdrA/NfnB fusions may represent a novel path back into the catabolic energy metabolism of ANME-1, confurcating electrons from NADPH and CoM-SH/CoB-SH.

The prevalence of many diverse copies of HdrABC systems in ANME-1 may offer an alternative explanation for how MetF-derived electrons are directed back into central catabolism. Recently a protein complex containing MetFV, HdrABC, and MvhD was discovered in *Moorella thermoacetica* that is suggested to carry out electron bifurcation with NADH, CH_2=H4F and other unknown reactants (Mock et al., 2014). Genomic analysis revealed the genes coding for these six proteins were found in a gene cluster together in *M. thermoacetica* and our examination of this HdrA homolog in this complex is much like HdrA 3 and 4 described above, with two fused HdrA domains with complementary Cys197/N-terminal domain loss. While none of the ANME-1 genomes analyzed here contain MetF in a gene cluster with HdrA homologs, the two "*Ca.* Syntrophoarchaeum" genomes both contain gene clusters containing MetFV, HdrA, and MvhD. If MetF catalyzes the CH_3-H_4MPT oxidation step in ANME-1, it may form complexes with one of the many HdrA paralogs. This would allow electrons to funnel directly into a bifurcation/confurcation reaction without passing through the $F_{420}H_2$ pool. In any case, electron flow through ANME-1 is likely to involve a significant amount of novelty that will require detailed biochemical studies to test these genomic predictions.

An energetic argument for both chemical diffusion and direct electron transfer in ANME-SRB syntrophy

Two strategies for syntrophic electron transfer could be widespread in the ANME-SRB syntrophy based on the conserved features of energy metabolism described here, EET utilizing the abundant MHC proteins, as well as diffusion of soluble electron carriers potentially produced by the diverse HdrABC complexes. A possible explanation for the conservation of these two systems in the genomes of ANME is that a mixed electron transfer through both systems may best suit an equal division of energy between the ANME and their sulfate reducing syntrophic partners.

If ANME-2a were to use Fpo, Rnf and HdrDE to oxidize $F_{420}H_2$, Fd^{2-}, and CoM-SH/CoB-SH, our expectation would be that all eight electrons from methane oxidation would end up on MpH_2 (**Figure 19A**). This would conserve abundant energy for the ANME through the Fpo and Rnf complexes, but the redox potential of the Mp/MpH_2 couple (E_0' = -165mV) is more than 50mV more positive than the SO_4^{2-}/HS^- redox couple (E_0' = -217mV). If all electrons were to pass through MpH_2 to the SRB their entire metabolism would be endergonic by about 40kJ/mol sulfate. For an equitable sharing of energy between ANME and SRB, the average redox potential of electrons during transfer should be approximately halfway between the CH_4/CO_2 (E_0' = -240mV) redox potential and the SO_4^{2-}/HS^- couple (E_0' = -217mV). The MpH_2 pool would need to be more than 90% reduced to drop its midpoint potential to ~ -230mV from -165mV, but this would render the HdrDE reaction impossible to run in the CoM-SH/CoB-SH oxidizing direction since the heterodisulfide E_0' is approximately 100mV more positive than this. The situation would be even more difficult for ANME-1 which are predicted to use a menaquinone which is very likely at a higher potential than MpH_2, requiring the membrane-bound electron carrier pool to be greater than 99% reduced for the SRB to yield any energy accepting all eight

electrons from menaquinone. In addition, whatever the MpH_2 or QH_2 pool midpoint potential is, it will be lower than the potential the SRB will receive electrons at, since the transfer from ANME to SRB through MHC must be down a redox gradient to maintain electron flow.

On the other hand, if all methane-derived electrons were used to produce a low-potential soluble electron carrier like formate through the action of cytoplasmic HdrA-based confurcation then the SRB could readily grow with electrons at this potential for sulfate reduction. In this scenario however the ANME's entire energy metabolism would contain no steps that could be coupled to energy conservation. Energy would be lost at the sodium pumping Mtr step, and there is no mechanism for recovering this energy since these soluble Hdr/Fdh complexes do not produce transmembrane potential or carry out substrate-level phosphorylation (**Figure 19B**).

A combined model of cytochrome-based EET and diffusion of low potential intermediates could alleviate the need for extremely reduced quinol pools to overcome the energetic inequalities between ANME and SRB, and may help explain the conservation and high expression of these two seemingly opposing methods of syntrophic electron transfer (**Figure 19C**). Four electrons transferred from two $F_{420}H_2$ through MpH_2/QH_2 and MHC-based EET could provide just enough energy to the ANME to overcome the losses at Mtr and support growth. This would utilize the Fpo and Fqo homologs that are universally conserved in all ANME groups. If the remaining four electrons are confurcated to formate or a similarly low potential electron carrier then the effective midpoint potential of this compound could be tuned to reach a balance where the effective redox potential of all eight of the transferred electrons were ~ -230mV, with the ANME conserving energy via the four electrons passing through Fpo/Fqo, and the SRB conserving energy through the four electrons on the lower potential soluble electron carrier.

The SRB would need to continue consuming the higher potential EET electrons, or else the lower potential electrons would cease flowing. Experimental results argue against formate as the predominant free extracellular intermediate in AOM (Meulepas et al., 2010; Nauhaus et al., 2005, 2002; Wegener et al., 2016), but perhaps another small molecule or diffusible organic substrate of sufficiently low redox potential could solve this energetic puzzle.

Conclusion

In proposing that AOM might be explained by methanogens running in reverse and transferring intermediates to a sulfate reducing bacterial partner, Zehnder and Brock concluded (Zehnder and Brock, 1980):

> "*We do not pretend that our hypothesis provides a complete explanation for the geochemical observations in anoxic marine environments, but our results provide reasonable evidence that methane is not biologically inert in strict anaerobic ecosystems.*"

In a similar spirit, we present these genome-based models of ANME metabolism with the understanding that they represent a series of hypotheses that require biochemical, genetic, or physiological experiments performed on purified ANME or ANME-encoded proteins to test. These models in the context of our expanded phylogenetic framework will help address important remaining questions in ANME-SRB syntrophy. Future studies will need to focus on specific ANME subgroups, preferably at the species level, to address questions of energy conservation mechanisms and adaptation to specific niches in anoxic marine environments. Novel methods of analysis such as electrode cultivation in may provide important new insights into the genomically predicted EET-capabilities. Features specifically supporting

syntrophy, including extracellular structures such as pili, specific interactions using phage-like protein translocation structures, or dockerin-cohesin systems will require careful experimental analysis far beyond genomic studies, but may have significant impacts on our understanding of cell-cell interactions in AOM consortia. We expect that explicit consideration of the carbon and electron transfer pathways in specific ANME clades will help address controversial results such as net methane production and the question autotrophic growth. Solving the enigma of the origin and functioning of the anaerobic oxidation of methane on Earth is important for understanding of the dynamics of methane fluxes in Earth history, and also for the evolution of the archaea and their syntrophic interactions with bacteria.

Materials and Methods

Quality and taxonomic assessment of genomes

Details of sample processing, DNA extraction, amplification (when applicable), sequencing and genome assembly for each genome can be found in **Supplementary file 4**. All genomes were assessed for their completeness and contamination using checkM 1.0.6 (Parks et al., 2015) using the 'taxonomy_wf' command and the Euryarchaeota marker set containing 188 marker genes. Genome taxonomy was assigned using the GTDB_Tk (Chaumeil et al., 2020) and naming conventions are consistent with the GTDB release 89.

16S rRNA gene sequencing of sorted single aggregates

For sorted single aggregate samples that did not have a 16S rRNA gene in the binned genomic dataset, we attempted to amplify the 16S rRNA genes for cloning and subsequent sequencing. The original MDA amplified samples were used as template to amplify the bacterial 16S rRNA gene using a modified universal reverse primer u1492R (5'-GGYTACCTTGTTACGACTT-3') and modified bacterial b8F-ym (5'-AGAGTTTGATYMTGGCTC- 3') (Lane, 1991). The u1492R primer was also paired with a

modified archaeal a8F-y (5'-TCCGGTTGATCCYGCC- 3') to amplify the archaeal 16S rRNA gene (von Wintzingerode et al., 1999). Fifteen microliter PCRs were run for 1 min at 94°C, followed by 40 cycles of 15s at 94°C, 30s at 54°C, and 45s at 72°C, and a final elongation step at 72°C for 4 min. The reaction mix included 0.2 μM forward and reverse primers (Integrated DNA Technologies, Inc. Coralville, IA, USA), 1× ExTaq PCR buffer (Takara, Clontech Laboratories, Inc., Mountain View, CA, USA), 0.75 U ExTaq (Takara), 0.22 mM dNTPs (New England Biolabs, Ipswich, MA, USA), and 1 μl template. No amplicon was visible in triplicate DNA control reactions when 4 μl were run on a 1.5% agarose gel visualized with SYBR Safe (ThermoFisher Scientific; cat. no. S33102). PCR products were plate purified (Millipore Multiscreen filter plates; cat. no. MSNU03010), ligated with the Invitrogen TOPO TA Cloning Kit (ThermoFisher Scientific; cat. No. K457501) and transformed using Top Ten chemically competent Escherichia coli cells. Picked colonies were grown overnight in Luria–Bertani broth (Miller's modification) and amplified using M13 primers for 30 cycles. The M13 products were visualized to confirm the correct size insert and screened by RFLP digest using HaeIII enzyme (NEB) on a 3% agarose gel. All unique clones were plate purified and sent for sequencing using T3 and T7 primers at Laragen Sequencing (Culver City, CA, USA). Contigs were assembled using Geneious software version R10 (Kearse et al., 2012). Sequence accessions for the clones from these genomes can be found in **Supplementary file 5**.

Calculation of genome similarity

The average amino acid identity (AAI) was calculated between each pair of genomes was calculated using compareM and its AAI workflow (aai_wf) command (https://github.com/dparks1134/CompareM). Briefly this workflow first identifies all open reading frames were determined using prodigal 2.0.6 (Hyatt et al., 2010) and then homologous genes were determined using reciprocal best-hits using 0.7.11

(Buchfink et al., 2015). The AAI was calculated by averaging the percent similarity of all proteins that were reciprocal best hits. The genome similarity (GS) displayed in **Table 1—figure supplement 1** was calculated by first calculating the percent alignment between the genomes as the fraction of the genes that are reciprocal best hits (RBH), using the smaller number of total genes between the two genomes (X & Y) as the denominator, and then multiplying the AAI value by that fraction (i.e. GS = AAI*RBH/min(X, Y)). The order of the matrix was calculated using the R statistical programming language 3.3.2 (R Core Team 2016) using the 'hclust' function using the default distance metric ('euclidean') and the 'ward.D2' method of agglomeration. This matrix was visualized using ggplot2 (Wickham, 2009).

16S rRNA gene phylogeny

16S rRNA gene sequences were extracted from SILVA 128 (Quast et al., 2013; Yilmaz et al., 2014) and new sequences from ANME genomes were aligned using SINA 1.2.11 (Pruesse et al., 2012) though the online portal at https://www.arb-silva.de/aligner/. Alignment columns were masked in ARB 6.0.2 (Ludwig et al., 2004) using the inbuilt 'archaea ssuref' filter and exported in fasta format removing all columns that contained only gaps. The phylogenetic tree in **Figure 1A** was constructed using RAxML 8.1.7 (Stamatakis, 2014) using the following parameters: raxmlHPC-PTHREADS-SSE3 -f a -k -x 12395 -p 48573 -N 100 -T 16 -m GTRGAMMAI. Trees were visualized using ete3 (Huerta-Cepas et al., 2016).

Concatenated protein phylogeny

The genome trees in **Figures 1B, 3** and **18** were constructed from a list of 43 hidden markov models (HMM) trained from proteins common to both bacteria and archaea sourced from Pfam and TIGRfam (**Supplementary file 6**). Open reading frames were annotated onto genome sequences using prodigal

2.6 (Hyatt et al., 2010) and candidates for the marker genes were identified using hmmsearch 3.1b1 (http://hmmer.org). Only the top match to each maker was retained and aligned to the HMM using hmmalign 3.1b1 (http://hmmer.org). The alignment was trimmed on each side using the stockholm format reference coordinate annotation (line beginning with #=GC RF) to include only residues that matched the reference (columns containing an 'x'). The protein sequences were then concatenated and a phylogenetic tree was constructed using RAxML 8.1.7 (Stamatakis, 2014) using the following parameters: raxmlHPC-PTHREADS-SSE3 -f a -k -x 67842 -p 568392 -N 100 -T 16 -m PROTGAMMAWAG -o IMG_638154518. Note that the outgroup is specified in the command line (-o IMG_638154518), which corresponds to Sulfolobus solfataricus.

McrA phylogeny

McrA sequences from cultured methanogenic and methanotrophic archaea and genome sequences from uncultured representatives were downloaded from the NCBI refseq database. These sequences were combined with clone sequences from ANME dominated environments (Lösekann et al., 2007; Trembath-Reichert et al., 2016) and from the ANME genomes sequenced here. These sequences were aligned using muscle 3.8.31 (Edgar, 2004) with the default parameters. The phylogenetic tree in **Figure 1D** was constructed using RAxML 8.1.7 (Stamatakis, 2014) using in the following parameters: raxmlHPC-PTHREADS-SSE3 -f a -k -x 67842 -p 19881103 -N 100 -T 16 -m PROTGAMMAWAG.

Transcriptome data

Data shown in **Supplementary file 2** was reproduced from the main text or supplemental data of papers reporting on the ANME-2a Wang genome (Wang et al., 2013), the ANME-2d BLZ1 genome (Arshad et al., 2015), and the ANME-2c E20, ANME-1 GB37 and GB60 genomes (Krukenberg et al., 2018).

NifD and FpoH phylogenies

Homologs of NifD and FpoH were identified in ANME genomes using KEGG ortholog IDs K02586 and K00337, respectively. Genes annotated with these KEGG IDs were separately combined with a set of protein sequences from uniprot that were also annotated with these IDs using the following website (in the case of NifD (http://www.uniprot.org/uniprot/? query=k02586&sort=score&format=fasta). A representative set of genes was made using cd-hit 4.6 (Li and Godzik, 2006) by removing all sequences that were greater than 95% amino acid similarity to each other. In the case of FpoH most bacterial sequences were omitted. These sequences were aligned using muscle 3.8.31 (Edgar, 2004) with the default parameters. A phylogenetic tree was constructed using RAxML 8.1.7 (Stamatakis, 2014) using in the following parameters: raxmlHPC-PTHREADS-SSE3 -f a -k -x 8512 -p 110339 -N 100 -T 16 -m PROTGAMMAWAG. Tree was visualized using ete3 (Huerta-Cepas et al., 2016).

Additional protein phylogenies

Phylogenetic trees for RpoB, MtrE, Mer, Mtd, Mch, Ftr, FwdB/FmdB, MetF, HdrA, VhtC/McoA, Small ANME MHCs, CcmF, GlyA, ThyA, FrhB were constructed based on alignments using muscle 3.8.31 (Edgar, 2004) with the default parameters. Trees were built with PhyML 3.0 (Guindon et al., 2010) implemented on the website (http://www.atgc-montpellier.fr/phyml/) with automatic substitution model selection based on the Akaike Information Criterion as implemented by SMS (Lefort et al., 2017). Branch support was provided by the aLRT SH-like fast likelihood-based method. Trees were visualized with iTOL (Letunic and Bork, 2019).

Determination of cytochrome orthologs

Homologous proteins displayed in **Figure 9—figure supplement 3** were determined using proteinortho 5.16 (Lechner et al., 2011) using the option '-singles' to keep even unique proteins. Orthologs containing multiheme cytochromes were determined by counting the number of CxxCH motifs in the amino acid sequences. Orthologous groups where the median count of heme binding domains was greater than one were included for analysis. These groups were then curated to resolve any erroneous placement of individual proteins. Each group of cytochromes was aligned using muscle 3.8.31 (Edgar, 2004) and individual sequences that did not share conserved residues or contained unusual insertions or deletions were moved into separate orthologs. Proteins that were originally categorized as unique were also compared using blastp 2.6.0+ (Camacho et al., 2009) to determine if they could be placed into an orthologous group.

Analysis of ANME-MHCs

Cytochromes with greater than ten heme binding motifs were annotated using Interpro (Finn et al., 2017; Jones et al., 2014) and were categorized based on the presence of an S-layer domain, transmembrane helixes, and the presences and type of putative sugar binding/degrading motifs.

Identification of dockerin/cohesin domains

Proteins containing dockerin and/or cohesin domains were identified by predicting domains in all hypothetical proteins in all ANME genomes using the KBase platform's Annotate Domains app (v.1.0.7) with the "All domain libraries" function (drawing from COGs, CDD, SMART, PRK, Pfam, TIGRFAMs, and NCBIfam) (Arkin et al., 2018). All domain annotations were downloaded as csv files, and any protein containing a dockerin or cohesin domain was exported along with other domains therein to a separate csv file. These csv files can be found in **Supplementary file 3**.

Acknowledgments

We thank Stephanie A. Connon and Alice Michel for help sequencing 16S rRNA genes from some single aggregates. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported under Contract No. DE-AC02-05CH11231. We acknowledge the support of all crews of the expeditions mentioned here for sampling and work at sea, and the work of Susanne Menger for the cultivation of AOM consortia. Funding for this study was received by the DFG Leibniz grant and by the Max Planck Society to AB. Genomes of BONCAT-FACS sorted ANME-consortia were generated via a JGI Director Discretionary Project Award (to R.H. and V.J.O.). We thank Chief Scientist Susan Lang, the scientific party of the 2018 Lost City Expedition, and NSF support to Lang and Brazelton (OCE- 1536702/1536405)

Figures

Genome	Order	Family	Genus	Species	Type	Sample Location	Habitat	Publication	Accession	Comp	Cont	Size (Mb)	Scaffolds
ANME-3 sp. HMMV	Methanosarcinales	Methanosarcinaceae	Methanovorans	sp. nov. 1	MAG	Haakon Mosby Mud Volcano	submarine mud volcano, 1250 m water depth	**This study**	DQIP01000000	98.0	3	2.45	236
ANME-3 sp. HMMV2	Methanosarcinales	Methanosarcinaceae	Methanovorans	sp. nov. 2	MAG	Haakon Mosby Mud Volcano	submarine mud volcano, 1250 m water depth	**This study**	DQIO01000000	94.4	4	2.56	334
ANME-2c sp. COP1	Methanosarcinales	Methanogasteraceae	Methanogaster	sp. nov. 1	MAG	Coal Oil Point	natural hydrocarbon seep, 47m depth	Parks *et al.* 2017	DASC01000000	88.5	2	2.39	222
ANME-2c sp. Agg-C10	Methanosarcinales	Methanogasteraceae	Methanogaster	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QBUL01000000	45.7	1	1.53	277
ANME-2c sp. HMMV-459B1	Methanosarcinales	Methanogasteraceae	Methanogaster	sp. nov. 3	MAG	Haakon Mosby Mud Volcano	submarine mud volcano, 1250 m water depth	**This study**	QBUV01000000	94.7	2	1.69	226
ANME-2c sp. E20	Methanosarcinales	Methanogasteraceae	Methanogaster	sp. nov. 2	MAG	Elba	enrichment culture, 20ºC	Krukenberg *et al.* 2018	PQXF00000000	95.7	4	3.98	7
ANME-2c sp. AMVER4-21	Methanosarcinales	Methanogasteraceae	Methanogaster	sp. nov. 4	MAG	Amon Mud Volcano	submarine mud volcano, 1120 m water depth	**This study**	QBUT01000000	62.7	5	1.90	94
ANME-2c ERB4	Methanosarcinales	Methanogasteraceae	Methanogaster	sp. nov. 1	Fosmids	Eel River Basin, California	methane cold seep, 5ºC	**This study**	see Methods	81.6	229	10.13	172
ANME-2c AMVER4-31	Methanosarcinales	Methanogasteraceae	gen. nov. 1	sp. nov. 1	MAG	Amon Mud Volcano	submarine mud volcano, 1120 m water depth	**This study**	QBUS01000000	65.1	4	1.60	109
ANME-2c COP2	Methanosarcinales	Methanogasteraceae	gen. nov. 1	sp. nov. 1	MAG	Coal Oil Point	natural hydrocarbon seep, 47m depth	Parks *et al.* 2017	DAWR01000000	92.5	2	2.37	224
ANME-2a S7142MS1	Methanosarcinales	Methanocomedenaceae	Methanocomedens	sp. nov. 1	MAG	Santa Monica Mounds, California	cold seep sediment incuabtion, 6ºC	Yu *et al* 2018	PYCK01000000	94.5	0	1.89	267
ANME-2a CONS7142H05b1	Methanosarcinales	Methanocomedenaceae	Methanocomedens	sp. nov. 1	ASAG	Santa Monica Mounds, California	cold seep sediment incuabtion, 6ºC	**This study**	QGHJ01000000	62.3	1	1.72	166
ANME-2a CONS7142G09b1	Methanosarcinales	Methanocomedenaceae	Methanocomedens	sp. nov. 1	ASAG	Santa Monica Mounds, California	cold seep sediment incuabtion, 6ºC	**This study**	QCYW01000000	70.3	0	2.14	193
ANME-2a HMMV-459B4	Methanosarcinales	Methanocomedenaceae	Methanocomedens	sp. nov. 1	MAG	Haakon Mosby Mud Volcano	submarine mud volcano, 1250 m water depth	**This study**	QBUQ01000000	83.2	4	1.45	501
ANME-2a Wang	Methanosarcinales	Methanocomedenaceae	Methanocomedens	sp. nov. 1	ASAG	Gulf of Cadiz	submarine mud volcano, 1200 m water depth	Wang *et al.* 2013	IMG-2565956544	98.5	3	3.64	575
ANME-2a HMMV-459A3	Methanosarcinales	Methanocomedenaceae	Methanocomedens	sp. nov. 2	MAG	Haakon Mosby Mud Volcano	submarine mud volcano, 1250 m water depth	**This study**	QBUR01000000	70.9	0	1.17	192
ANME-2b HR1	Methanosarcinales	Methanocomedenaceae	Methanomarinus	sp. nov. 1	MAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	Yu *et al* 2018	MZXQ01000000	95.7	0	2.20	335
ANME-2b CONS3730F09p3b1	Methanosarcinales	Methanocomedenaceae	Methanomarinus	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QCYX01000000	79.1	0	2.03	213
ANME-2b CONS3730E01UFb2	Methanosarcinales	Methanocomedenaceae	Methanomarinus	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QENI01000000	71.8	1	2.14	224
ANME-2b CONS3730D10UFb2	Methanosarcinales	Methanocomedenaceae	Methanomarinus	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QYBA01000000	58.9	0	1.42	266
ANME-2d *Ca*. Methanoperedens sp. BLZ1	Methanosarcinales	Methanoperedenaceae	Methanoperedens	nitroreducens	MAG	Wastewater bioreactor	bioreactor	Arshad *et al.* 2015	LKCM01000000	96.8	3	3.74	514
ANME-2d *Ca*. Methanoperedens nitroreducens	Methanosarcinales	Methanoperedenaceae	Methanoperedens	nitroreducens	MAG	Wastewater bioreactor and sediment	bioreactor	Haroon *et al.* 2013	JMIY01000000	98.8	1	3.20	10
ANME-1 AG-394-G06	Methanophagales	Methanophagaceae	Methanophaga	sp. nov. 1	ASAG	Gulf of Mexico	methane cold seep, 5ºC	**This study**	PRCZ01000000	66.0	1	1.53	119
ANME-1 AG-394-G21	Methanophagales	Methanophagaceae	Methanophaga	sp. nov. 1	ASAG	Gulf of Mexico	methane cold seep, 5ºC	**This study**	PRCY01000000	68.5	1	1.87	101
ANME-1 Meyerdierks	Methanophagales	Methanophagaceae	Methanophaga	sp. nov. 1	MAG	North-Western Crimean sea	microbial mat, ~200 m water depth	Meyerdierks *et al.* 2010	FP565147	82.3	21	3.40	1
ANME-1 GoMg3.2	Methanophagales	Methanophagaceae	Methanophaga	sp. nov. 1	Co-SAG	Gulf of Mexico	methane cold seep, 5ºC	**This study**	QBUN01000000	81.5	6	2.73	426
ANME-1 ERB7	Methanophagales	Methanophagaceae	Methanophaga	sp. nov. 1	Fosmids	Eel River Basin, California	methane cold seep, 5ºC	**This study**	see Methods	90.7	247	5.87	175
ANME-1 CONS3730B06UFb1	Methanophagales	Methanophagaceae	Methanophaga	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QENH01000000	90.1	2	3.18	201
ANME-1 Agg-C03	Methanophagales	Methanophagaceae	gen. nov. 1	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QBUP01000000	65.4	3	1.90	375
ANME-1 CONS3730F07p2b1	Methanophagales	Methanophagaceae	gen. nov. 1	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QENK01000000	74.6	3	2.75	380
ANME-1 CONS3730MDAH03UFb1	Methanophagales	Methanophagaceae	gen. nov. 1	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QENL01000000	79.1	15	3.42	505
ANME-1 CONS3730H04p2b1	Methanophagales	Methanophagaceae	gen. nov. 1	sp. nov. 1	ASAG	Hydrate Ridge, Oregon	cold seep sediment incuabtion, 6ºC	**This study**	QENJ01000000	64.4	1	2.57	369
ANME-1 GoMg2	Methanophagales	Methanophagaceae	gen. nov. 1	sp. nov. 2	Co-SAG	Gulf of Mexico	methane cold seep, 5ºC	**This study**	QBUM01000000	79.5	3	2.19	251
ANME-1 GoMg4	Methanophagales	Methanophagaceae	Methanoalium	sp. nov. 1	Co-SAG	Gulf of Mexico	methane cold seep, 5ºC	**This study**	QBUO01000000	55.9	2	1.34	177
ANME-1 SA	Methanophagales	Methanophagaceae	Methanoalium	sp. nov. 2	MAG	Welkom, South Africa	Gold Mine Aquifer, 37ºC	Lau *et al* 2014	mgm4536100.3	91.3	2	2.01	196
ANME-1 ERB6	Methanophagales	Methanophagaceae	gen. nov. 2	sp. nov. 1	Fosmids	Eel River Basin, California	methane cold seep, 5ºC	**This study**	see Methods	66.4	82	4.55	134
ANME-1 SCGC AAA252-L18	Methanophagales	Methanophagaceae	gen. nov. 2	sp. nov. 1	Co-SAG	Gulf of Cadiz	submarine mud volcano, 3860 m water depth	**This study**	DQIN00000000	71.4	6	2.57	700
ANME-1 GoMg1	Methanophagales	Methanophagaceae	gen. nov. 2	sp. nov. 1	Co-SAG	Gulf of Mexico	methane cold seep, 5ºC	**This study**	QBUL01000000	80.5	4	2.34	189
ANME-1 ex4572-4	Methanophagales	Methanophagaceae	gen. nov. 3	sp. nov. 1	MAG	Guaymas basin	orange microbial mat, 50-80ºC	Dombrowski *et al.* 2017	NJEM01000000	70.9	2	1.02	95
ANME-1 GB37	Methanophagales	Methanophagaceae	gen. nov. 3	sp. nov. 1	MAG	Guaymas basin	enrichment culture, 37ºC	Krukenberg *et al.* 2018	PQXB01000000	92.1	4	1.53	4
ANME-1 GB60	Methanophagales	Methanophagaceae	gen. nov. 4	sp. nov. 1	MAG	Guaymas basin	enrichment culture, 60ºC	Krukenberg *et al.* 2018	PQXC00000000	88.9	2	1.62	36

Table 1 Genome statistics Summary of genome characteristics, sources and accession numbers of the analyzed ANME genomes. Major ANME subgroups are separated by bold lines and proposed taxonomy adheres to the GTDB framework. At least one genus within each subgroup has been given a proposed name, and genomes that fall within other genera or species are simply given numerical placeholders (Genus and Species columns, respectively). Type indicates the method used for genome reconstruction, including: metagenome assembled genome (MAG), amplified single aggregate genome (ASAG), combine-assembled single amplified genomes (Co-SAGs), or fosmid library. Estimated levels of completeness (comp) and contamination (cont) are reported as percent as evaluated by CheckM. Note: ERB4, ERB6, and ERB7 are fosmid libraries with highly similar strains combined which has resulted in high levels of "contamination". See **Supplementary file 1** for complete details of assembly and binning for each genome.

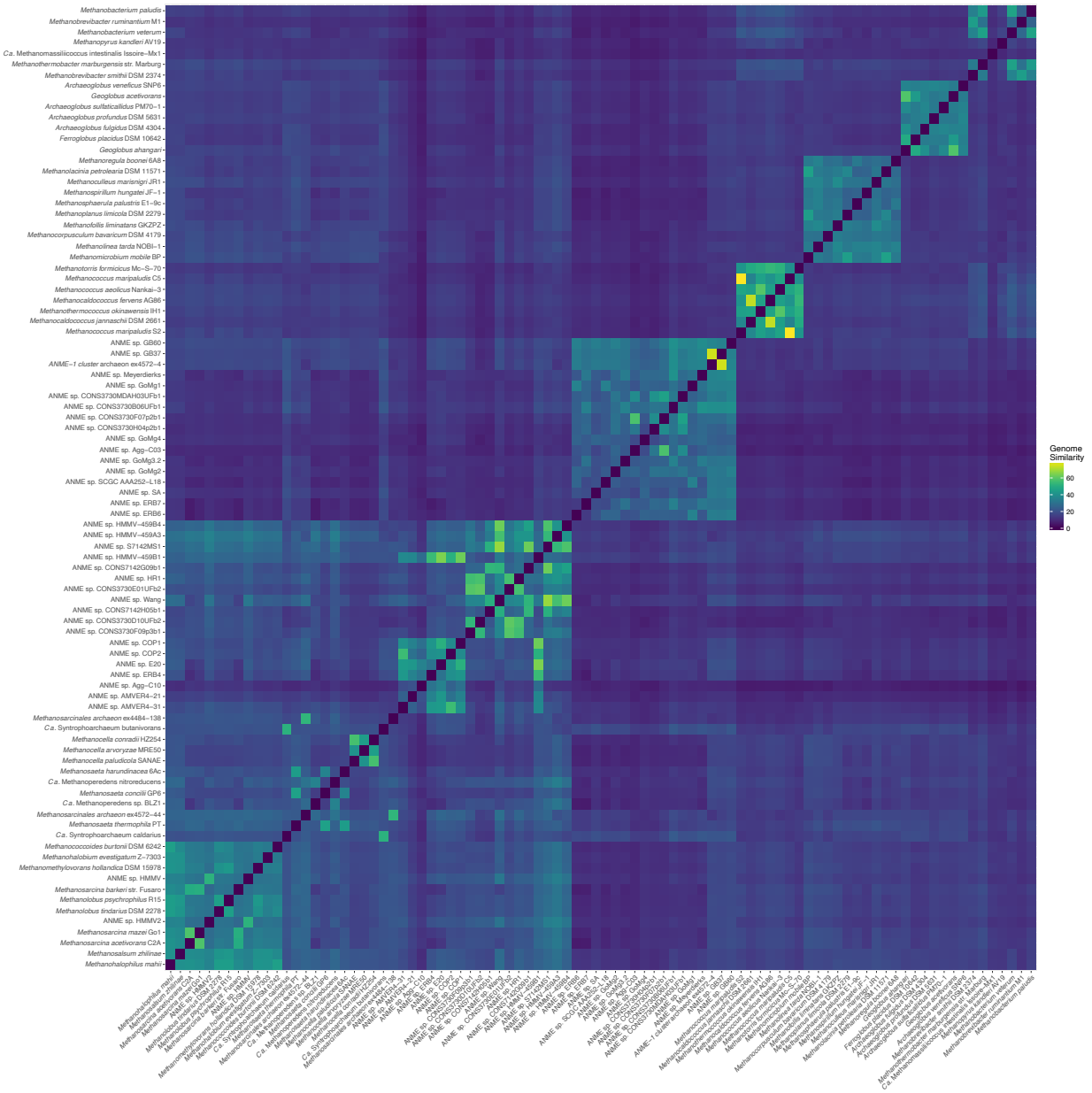


Table 1—figure supplement 1 Genome similarity between methanogen and ANME genomes reported here. Details of similarity calculations can be found in the Materials and Methods.

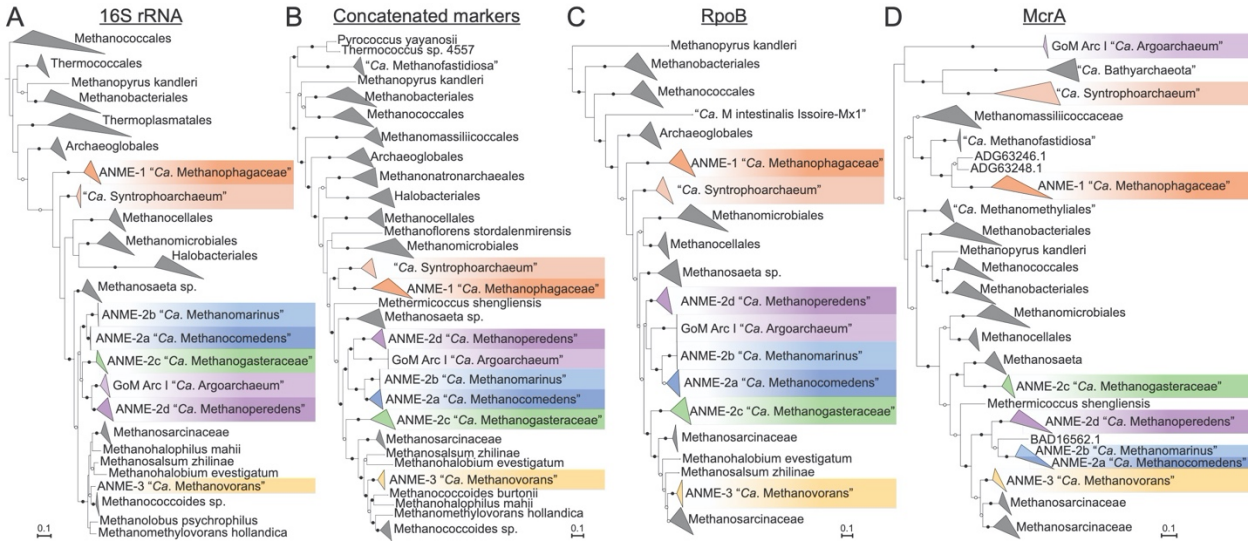


Figure 1 Phylogeny of ANME and related archaea Phylogenetic trees constructed from ANME genomes, sequences of closely related archaea, and a selection of sequences derived from clone libraries demonstrate the relationship between ANME and methanogens of the *Halobacterota*. (**A**) Phylogenetic tree built with 16S rRNA gene sequences, root leads to sequence from *Sulfolobus solfataricus* p2. (**B**) Phylogenomic tree built with concatenated marker set 4 (see **Supplementary file 3** for list), root also to *S. solfataricus* p2. (**C**) Phylogenetic tree built with protein sequence of RpoB, root leads to sequences from "*Ca.* Methanomethyliales" and "*Ca.* Bathyarchaeota". (**D**) Phylogenetic tree of McrA protein sequences. Note the divergence of proposed alkane oxidizing McrA genes in "*Ca.* Syntrophoarchaeum", "*Ca.* Bathyarchaeota" and "*Ca.* Argoarchaeum". Branch support values of 100% are labelled with closed circles, >50% with open circles. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. For tree files and alignments see **Figure 1—source data 1**.

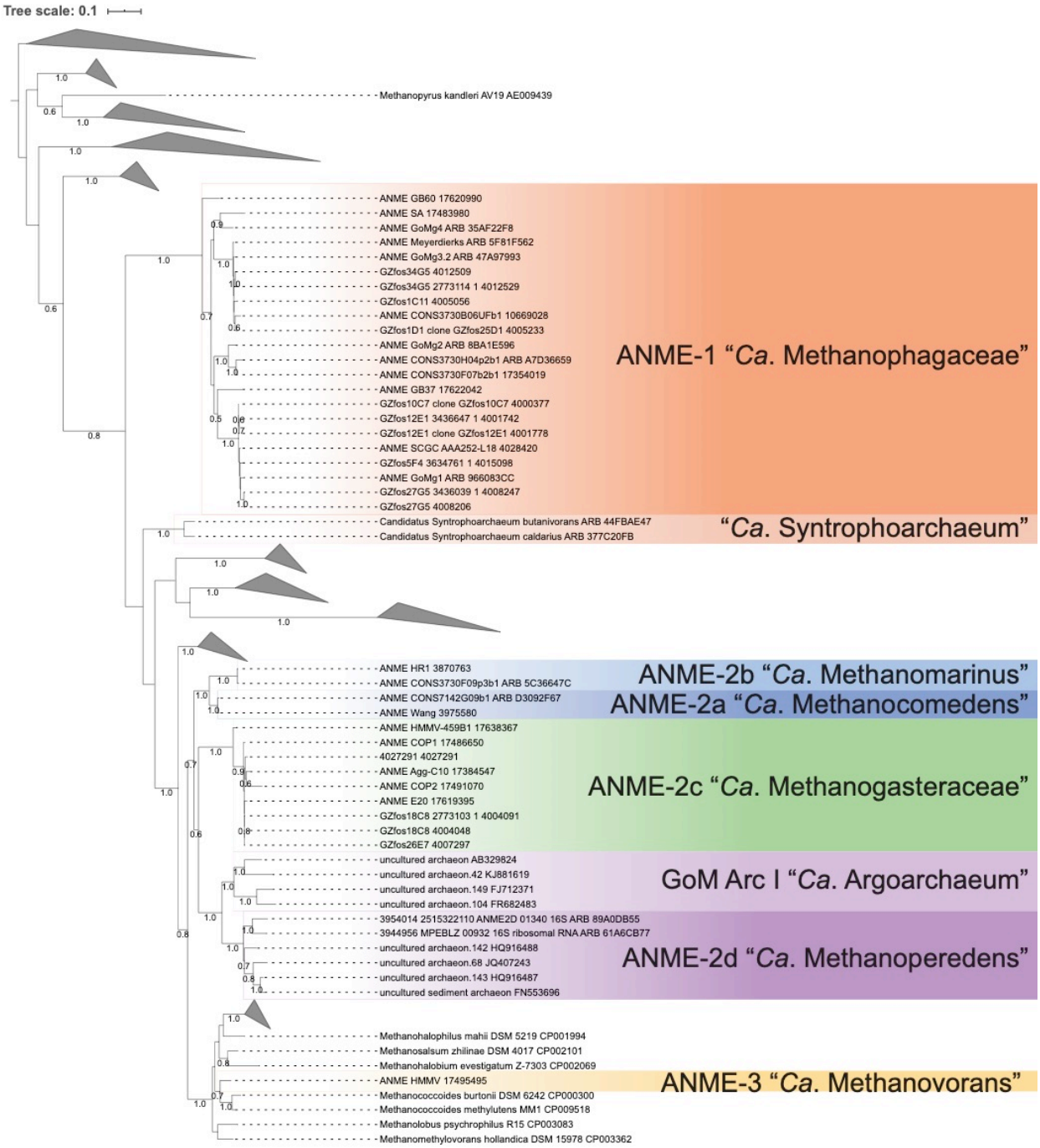


Figure 1—figure supplement 1 Expanded 16S rRNA gene tree
Tree expanding compressed clades in **Figure 1A**

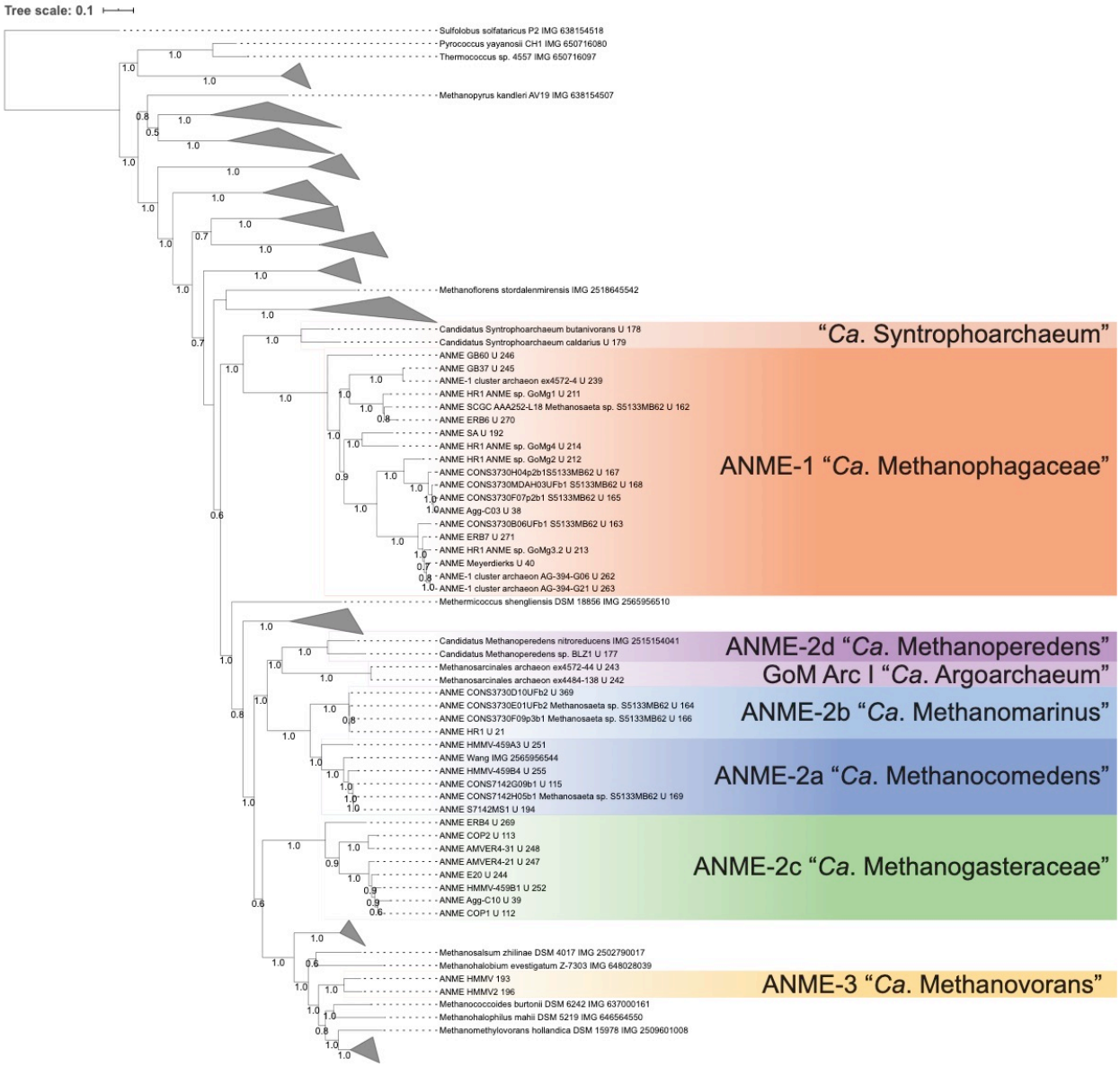


Figure 1—figure supplement 2 Expanded concatenated marker protein tree
Tree expanding compressed clades in **Figure 1B**

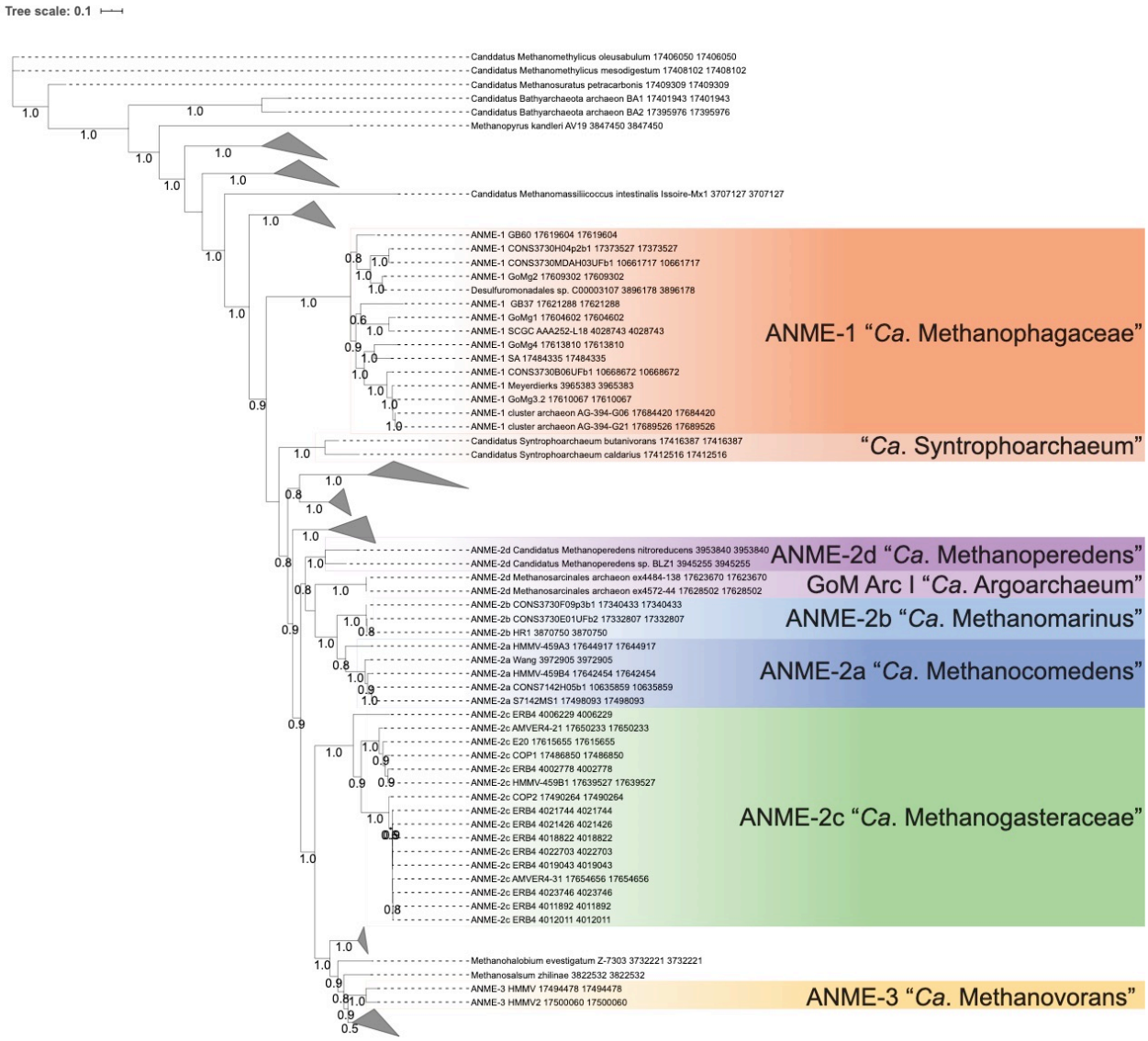


Figure 1—figure supplement 3 Expanded rpoB protein tree
Tree expanding compressed clades in **Figure 1C**

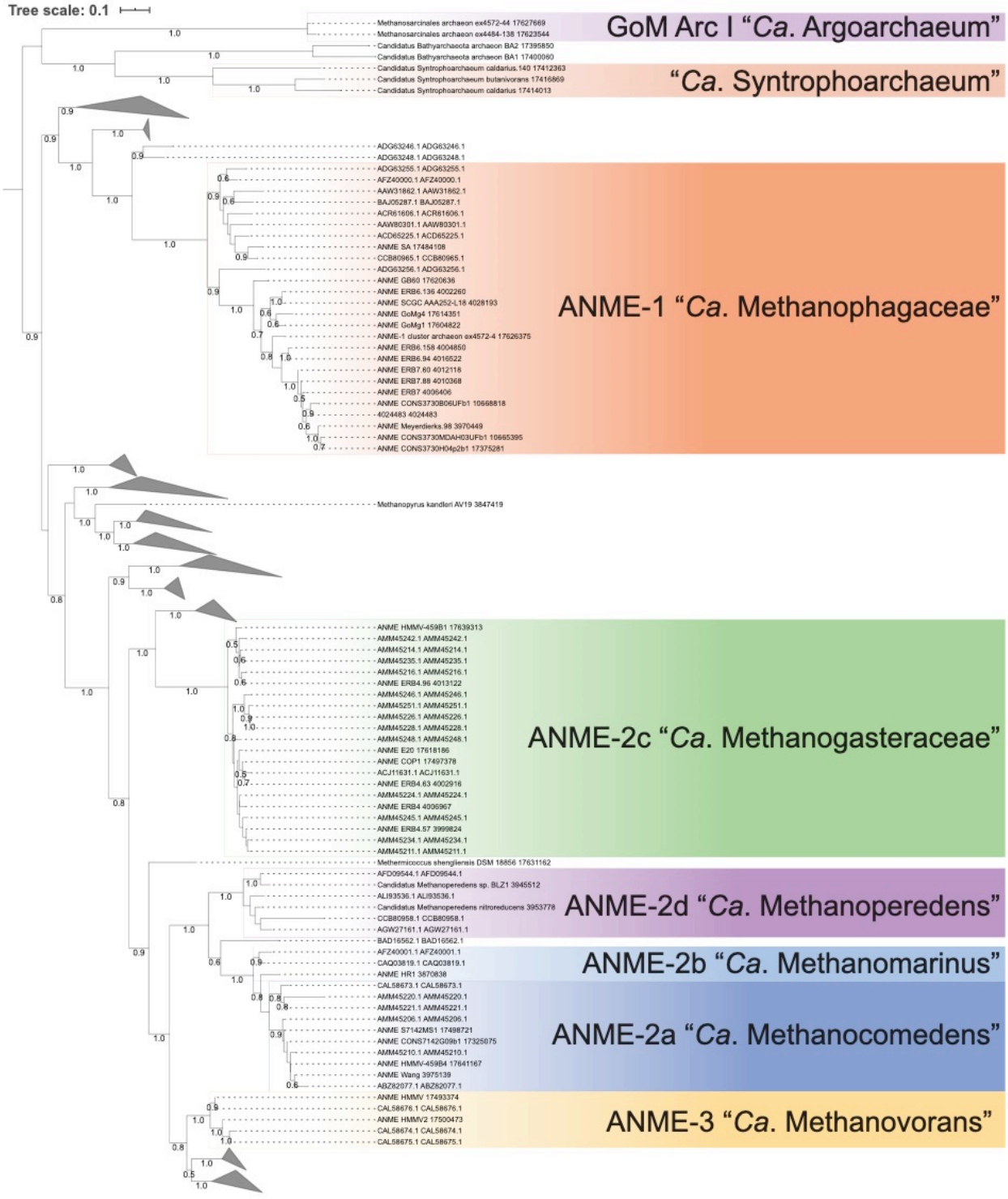


Figure 1—figure supplement 4 Expanded mcrA protein tree
Tree expanding compressed clades in **Figure 1D**

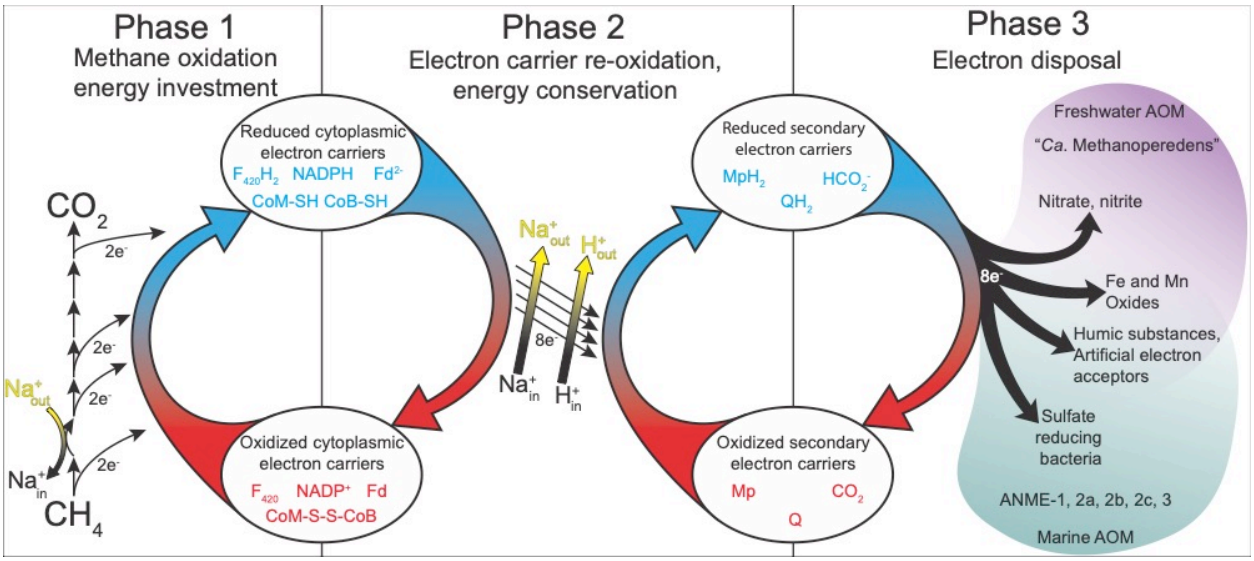


Figure 2 Summary of ANME energy metabolism Schematic representation of the three phases of ANME energy metabolism in our current model. In Phase 1 methane is oxidized to CO_2 through the reversal of the canonical seven step methanogenesis pathway. Energy is invested in this phase in the form of sodium ion translocation from the outer face of the cytoplasmic membrane to the inner face (yellow arrow). As C1 moieties are sequentially oxidized, eight electrons are transferred to soluble electron carriers such as $F_{420}H_2$, NADPH, Fd^{2-}, and CoM-SH/CoB-SH. In Phase 2 eight electrons on these primary electron carriers are transferred to secondary electron carriers in a process that conserves energy needed for cell growth in the form of sodium and proton motive forces (yellow arrows). In Phase 3 the secondary electron carriers are relieved of their electrons in various ways depending on the environmentally available electron acceptors, which can include sulfate reducing bacteria in the case of marine ANME-SRB consortia, iron, manganese or oxidized nitrogen species in the case of "*Ca.* Methanoperedens". Humic substances and artificial electron acceptors (AQDS) have also served as electron acceptors in laboratory experiments for a variety of different ANME groups from fresh and marine environments.

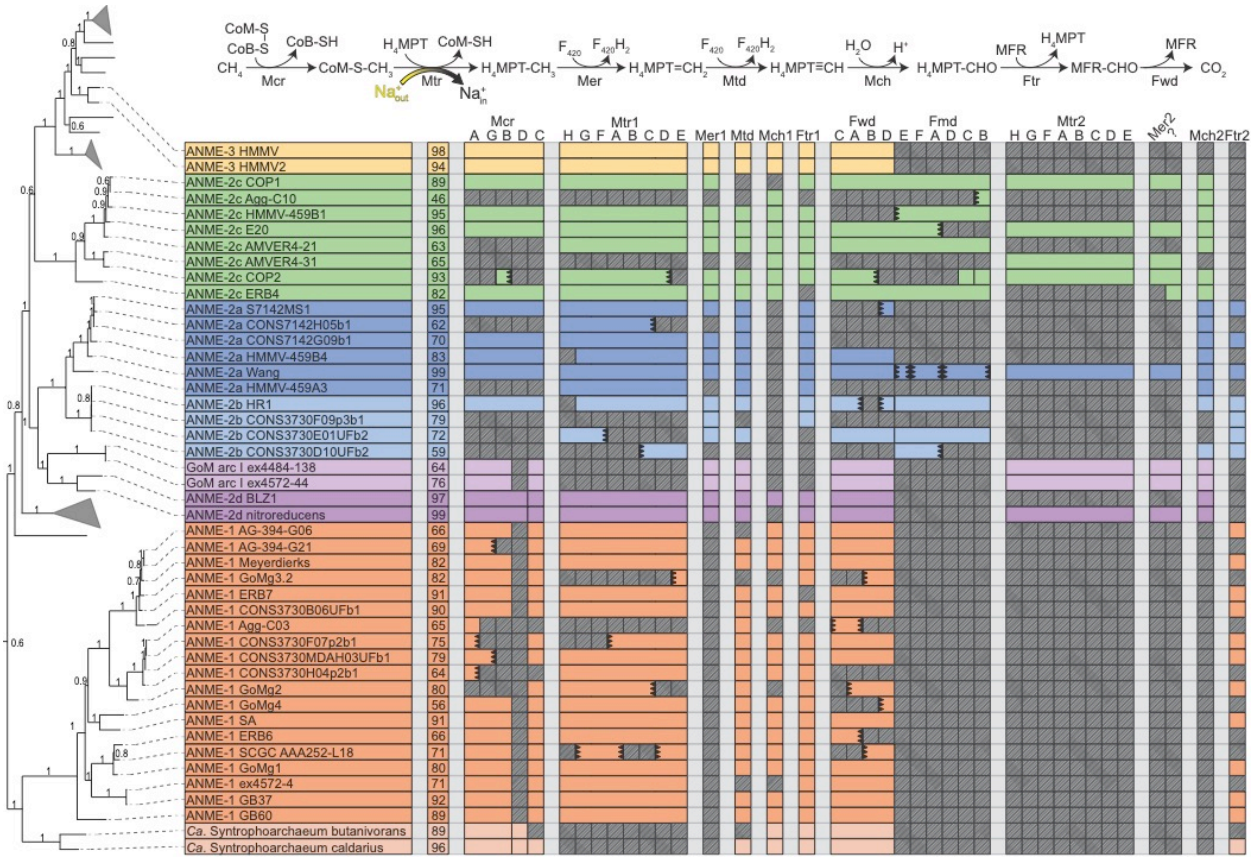


Figure 3 Presence of methanogenesis pathway genes in ANME archaea The proteins responsible for the seven steps of methanogenesis from CO_2. Mcr: Methyl-coenzyme M reductase; Mtr: N^5-methyl-H_4MPT:coenzyme M methyltransferase; Mer: methylene-H_4MPT reductase; Mtd: F_{420}-dependent methylene-H_4MPT dehydrogenase; Mch: N^5,N^{10}-methenyl-H_4MPT cyclohydrolase; Ftr: formylmethanofuran-H_4MPT formyltransferase; Fmd/Fwd: formyl-methanofuran dehydrogenase. Colored boxes represent presence of homologs of these proteins in ANME genomes. Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent estimated genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. If a gene cluster appears truncated by the end of a contig it is depicted by a serrated edge on the box representing the last gene on the contig. Numbers following protein names indicate whether the enzyme is closely related to those found in *Methanosarcinaceae* (1) or are distantly related homologs (2). Question mark represents hypothetical protein of unknown function found clustered with Mer2. Tree orienting genome order is the same as found **Figure 1B**. For details on paralog phylogenetic relations see **Figure 4**. Gene accession numbers can be found in **Figure 3—source data 1**.

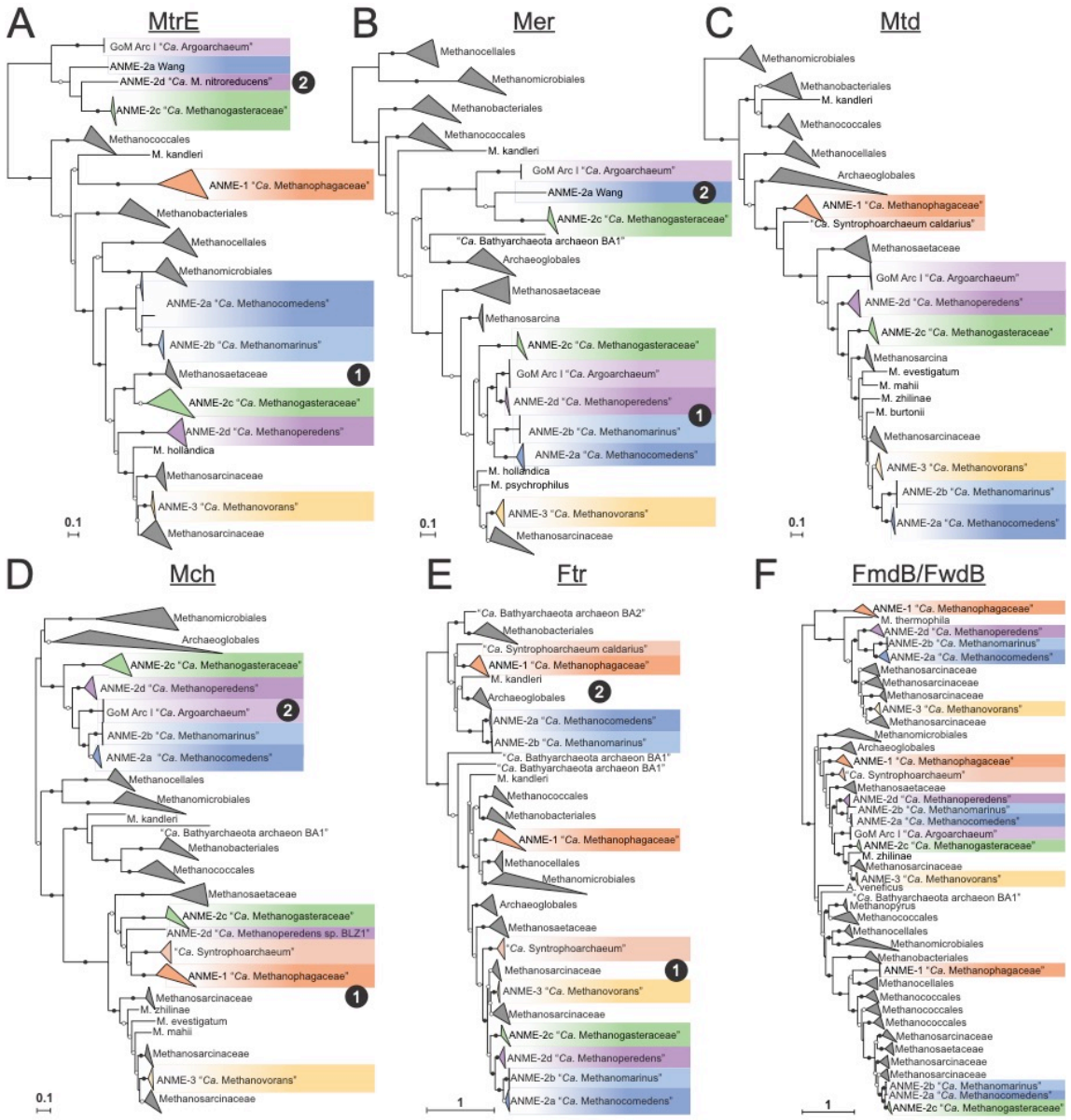


Figure 4 Methanogenesis phylogeny Phylogenetic trees constructed from protein sequences of enzymes involved in the methanogenesis pathway in ANME and related archaea. Mcr phylogeny is presented in **Figure 1**. Numbers next to clades indicates whether the cluster is closely related to those found in *Methanosarcinaceae* (1) or are distantly related homologs (2), matching labels in **Figure 3**. **A)** MtrE: N^5-methyl-H_4MPT:coenzyme M methyltransferase subunit E; **B)** Mer: methylene-H_4MPT reductase; **C)** Mtd: F_{420}-dependent methylene-H_4MPT dehydrogenase; **D)** Mch: N^5,N^{10}-methenyl-H_4MPT cyclohydrolase; **E)** Ftr: formylmethanofuran-H_4MPT formyltransferase; **F)** FmdB/FwdB: formyl-methanofuran dehydrogenase subunit B, molybdenum/tungsten variety, respectively. Branch support values of 100% are labelled with closed circles, >50% with open circles. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignments and tree files can be found in **Figure 4—source data 1**

	ANME-2a sp. Wang	ANME-2c sp. E20	ANME-1 sp. GB37	ANME-1 sp. GB60	ANME-2d sp. BLZ1
McrA	2409	10.1282	9.2004	11.3398	57436
McrB	4618	9.2748	9.1975	11.0564	31898
McrC	4112	9.2353	1.6110	4.1348	45848
McrD	3943	9.2396	x	x	NR
McrG	3112	9.6083	9.3888	11.4629	26015
MtrA1	1429	5.1720	4.6449	5.1880	1201
MtrB1	608	5.0568	5.0891	6.0365	1018
MtrC1	5359	4.7465	4.3046	5.5697	741
MtrD1	5080	5.0899	5.0314	5.7621	720
MtrE1	6057	4.8296	5.6337	6.1483	1268
MtrF1	2948	5.2516	4.3491	6.0343	2117
MtrG1	470	5.7992	4.9181	6.3578	1505
MtrH1	1720	5.6614	4.3311	7.0791	1915
MtrA2	230	0.7280	x	x	x
MtrB2	NM	1.0737	x	x	x
MtrC2	NM	0.7601	x	x	x
MtrD2	173	0.7152	x	x	x
MtrE2	5128	-0.3353	x	x	x
MtrF2	NM	2.0109	x	x	x
MtrG2	50		x	x	x
MtrH2	167	1.0137	x	x	x
Mer1	2652	7.7327	x	x	1575
Mer2	393	-2.7409	x	x	x
MetF	NR	x	3.3903	6.0426	NR
MetV	NR	x	2.7532	5.0344	NR
Mtd	66	7.1454	3.3944	3.5399	1130
Mch1	x	5.2326	0.9195	2.6930	374
Mch2	4872	3.6289	x	x	160
Ftr1	1581	5.3322	2.4222	3.7288	1333
Ftr2	NM	x	-1.9296	-0.0508	x
FwdA	6030	3.9056	2.4439	4.1120	404
FwdB	3030	3.7583	1.9024	4.4592	511
FwdC	7392	5.3057	2.8263	3.5015	691
FwdD	4536	4.0554	3.2972	3.9773	296
RnfA	622	3.4474	x	x	x
RnfB	427	3.7907	x	x	x
RnfC	1296	2.9705	x	x	x
RnfD	492	2.2104	x	x	x
RnfG	1012	2.2686	x	x	x
RnfE	650	2.6610	x	x	x
FpoA	2296	4.3301	1.6742	1.2893	304
FpoB	NM	3.5198	3.0775	2.7928	664
FpoC	NM	3.5301	x	x	393
FpoD	2160	4.1695	x	x	672
FpoH	1196	3.1499	2.2905	-	588
FpoI	2265	3.3646	3.1785	-	705
FpoJ1	NM	3.8686	0.7241	1.2708	645
FpoJ2	x	4.2977	x	x	479
FpoK	NM	4.8266	2.7599	3.0421	720
FpoL	2320	3.9066	1.6561	2.0428	752
FpoM	2484	3.7661	1.6597	2.2938	481
FpoN	3135	4.1567	1.9146	0.9632	618
FpoO	NM	x	x	x	730
FpoX	x	5.2315	x	x	x
FpoY	x	5.2654	x	x	x
FpoF	4368	5.3571	3.0433	3.9571	964
HdrD	2551	5.3612	x	x	946
HdrE	1349	4.5716	x	x	1369
HdrA1	1878	-0.7572	x	x	
HdrB1	1269	-1.0860	x	x	
HdrC1	1241	-0.4109	x	x	
HdrD1	287	-3.6243	x	x	
?	NR	3.7303	x	x	
FdhA1	NM	3.4938	x	x	
FrhB3	1326	2.4211	x	x	368
MvhD1	NR	2.7959	x	x	
HdrA2	NR	3.5522	x	x	
MvhD2	NM	4.1741	x	x	
HdrA3	1878	3.7475	x	x	
HdrB2	1269	2.9469	x	x	
HdrC2	760	3.2553	x	x	
HdrD2	NR	2.8987	x	x	
HdrA4	x	x	2.9911	x	
HdrB4	x	x	3.3477	x	
HdrC4	x	x	2.8278	x	
FrhB4	x	x	3.1032	x	
MvhD4	x	x	4.1699	x	
HdrA5 (11)	x	x	2.2274	4.7878	
HdrA4 (7)	x	x	2.6829	5.0229	
HdrB3	x	x	4.0340	5.5270	
HdrC3	x	x	3.9055	5.7005	
AmcA/VhtC	1539	5.3246	x	x	x
AmcB (cyt c)	NR	6.9548	x	x	x
Acd	685				

FrhB6	x	x	x	2.0495	
PorB (FrhB6)	x	x	x	1.5793	
PorA (FrhB6)	x	x	0.6903	2.0865	
PorD (FrhB6)	x	x	0.1959	2.9897	
PorG (FrhB6)	x	x	0.5254	2.5898	
FrhB7	x	x	0.4533	4.2476	
CODHe (FrhB7)	x	x	1.7939	x	
CODHa (FrhB7)	x	x	1.2966	2.2940	
FrhB8	NR	-1.6629	0.4619	1.8517	
FrhB8	NR	0.6158	1.4006	2.4513	
FrhB8	NR	-0.2244	1.9178	2.8206	
FrhB9 (Fsr)	NR	3.8021	1.6414	x	
FrhB10	x	x	x	-0.5359	
FrhB10	x	x	x	0.8471	

PLTS01	NR	1.4107	x	x	NR
PLTS02	NR	1.1459	x	x	NR
PLTS03	NR	0.9866	x	x	NR
PLTS04	NR	2.4941	x	x	NR
PLTS05	NR	3.0464	x	x	NR
PLTS06	NR	3.2592	x	x	NR
PLTS07	NR	2.5886	x	x	NR
PLTS08	NR	1.0300	x	x	NR
PLTS09	NR	0.6744	x	x	NR
PLTS10	NR	0.1041	x	x	NR
PLTS11	NR	0.3114	x	x	NR
PLTS12	NR	0.7103	x	x	NR
PLTS13	NR	1.6680	x	x	NR
PLTS14	NR	1.5961	x	x	NR
PLTS15	NR	1.3398	x	x	NR
PLTS16	NR	0.9900	x	x	NR
PLTS17	NR	0.6125	x	x	NR
PLTS18	NR	1.7670	x	x	NR

Highest MHC (from Fig 9-sup 4)	x	5.9957	7.0882	6.6859	9148
Highest MHC overall	x	7.2325	7.0882	6.7590	9148

x = not present in genome
NM = not measured
NR = not reported

Supplementary file 2 Summary of transcriptome results Multiple transcriptome studies have been conducted in conjunction with some of the ANME genomes previously published and reanalyzed here. The transcription level of specific ANME features highlighted in the present work have been compiled from the supplementary data of these studies.

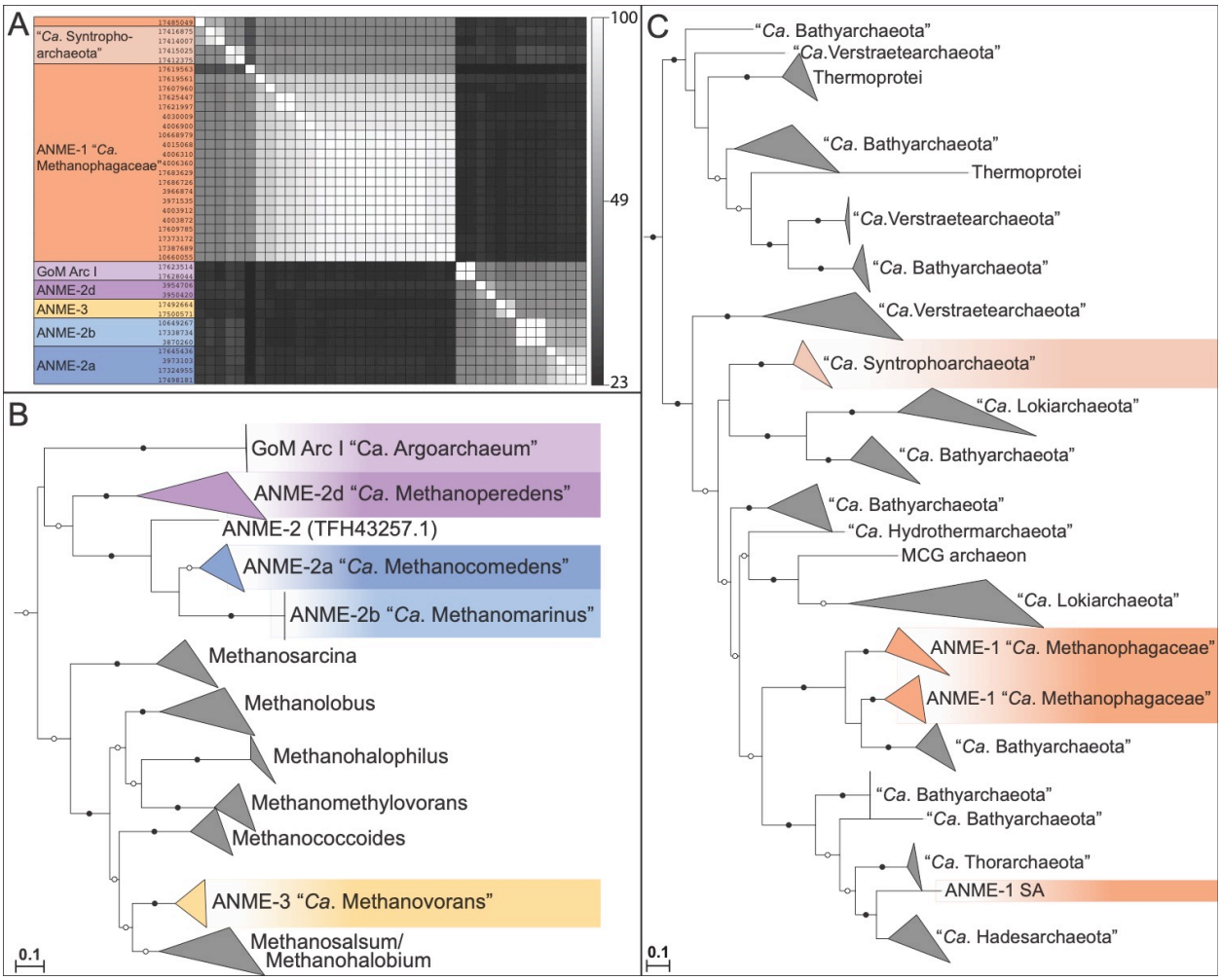


Figure 5 MetF in ANME Archaea (**A**) Amino acid sequence identity of MetF homologs found in ANME, "*Ca.* Argoarchaeum" and "*Ca.* Syntrophoarchaeum". ANME-1 and "*Ca.* Syntrophoarchaeum" form one cluster based on sequence similarity, while ANME-2a, 2b, 2d, 3 and "*Ca.* Argoarchaeum" form a second. Grayscale values represent percent identity. Sequences similar to the ANME-2/3 or ANME-1 clusters were retrieved via BLAST search of the NCBI nr database and used to construct phylogenetic trees of these two groups. (**B**) ANME-2, 3 and "*Ca.* Argoarchaeum" cluster together with closely related members of the *Methanosarcinaceae*. (**C**) ANME-1 and "*Ca.* Syntrophoarchaeum" form a polyphyletic group within a diverse group of sequences derived from MAGs of uncultured archaea. Notably the ANME-1 sp. SA is significantly different than the rest of the ANME-1. Roots for both trees lead to closely related MetF sequences from bacteria. Branch support values of 100% are labelled with closed circles, >50% with open circles. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignments and tree files can be found in **Figure 5—source data 1**.

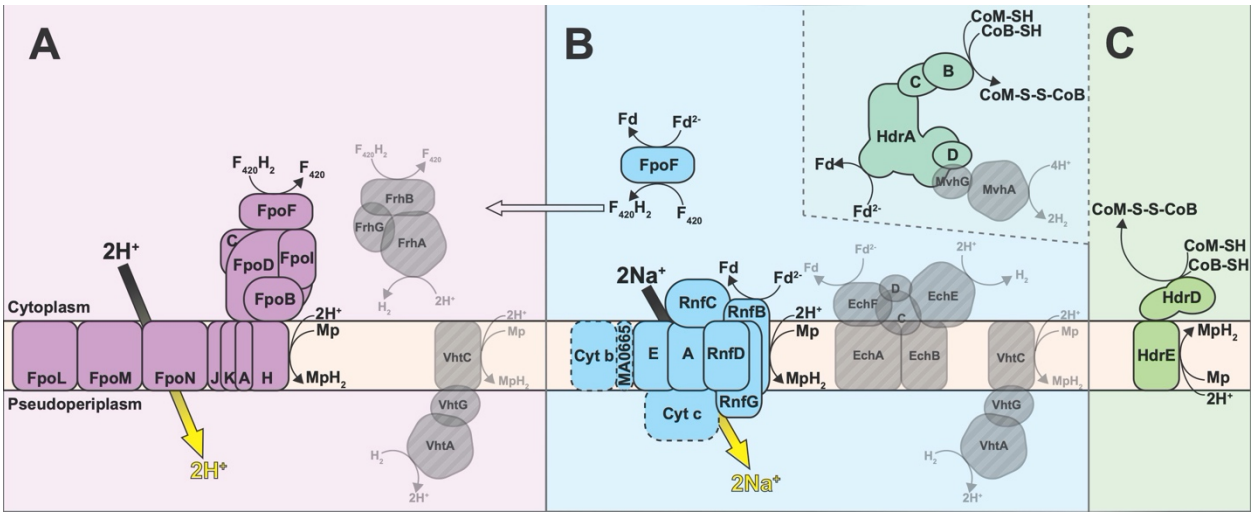


Figure 6 Cytoplasmic electron carrier oxidation Some energy conservation systems discovered in methanogenic archaea are conserved in ANME archaea (colored fill) while others appear absent (transparent gray with diagonal line fill). (**A**) $F_{420}H_2$ oxidation is coupled to proton translocation in methylotrophic methanogens via the Fpo/Fqo complex or by the production of H_2 by Frh and subsequent oxidation by Vht. In either case, electrons are ultimately deposited on MpH_2. Neither Frh or Vht complexes have been observed in any ANME genomes analyzed here. (**B**) Fd^{2-} oxidation can be coupled either to sodium motive force or proton motive force in methylotrophic methanogens. The Rnf complex catalyzes Fd^{2-}:Mp oxidoreductase reaction coupled to sodium translocation and is found in a number of methanogens and ANME. The ANME-2c contain most of the complex, but lack the cytochrome *c*, cytochrome *b* and MA0665 subunits, so their activity is difficult to predict. Ech and Vht can combine to produce net proton translocation via H_2 diffusion in methylotrophic methanogens, but neither complex is found in ANME. FpoF can catalyze a Fd^{2-}:F_{420} oxidoreductase reaction and $F_{420}H_2$ could then pass through the Fpo/Fqo complex. Various HdrABC complexes are present in all ANME genomes, and could in principle oxidize Fd^{2-} ad CoM-SH/CoB-SH through a reversal of electron bifurcation reaction. The electron acceptor in this process is likely to not be H_2 in most ANME groups due to the absence of MvhG and MvhA. (**C**) Besides the HdrABC complexes mentioned above and second possible CoM-SH/CoB-SH oxidation strategy would be a reversal of the HdrDE reaction found in methylotrophic methanogens. In ANME the reaction would have to proceed in the direction illustrated, and therefore would dissipate proton motive force by consuming a proton on the positive side of the membrane. For presence/absence of these systems in ANME genomes analyzed here see **Figure 6—figure supplement 1.**

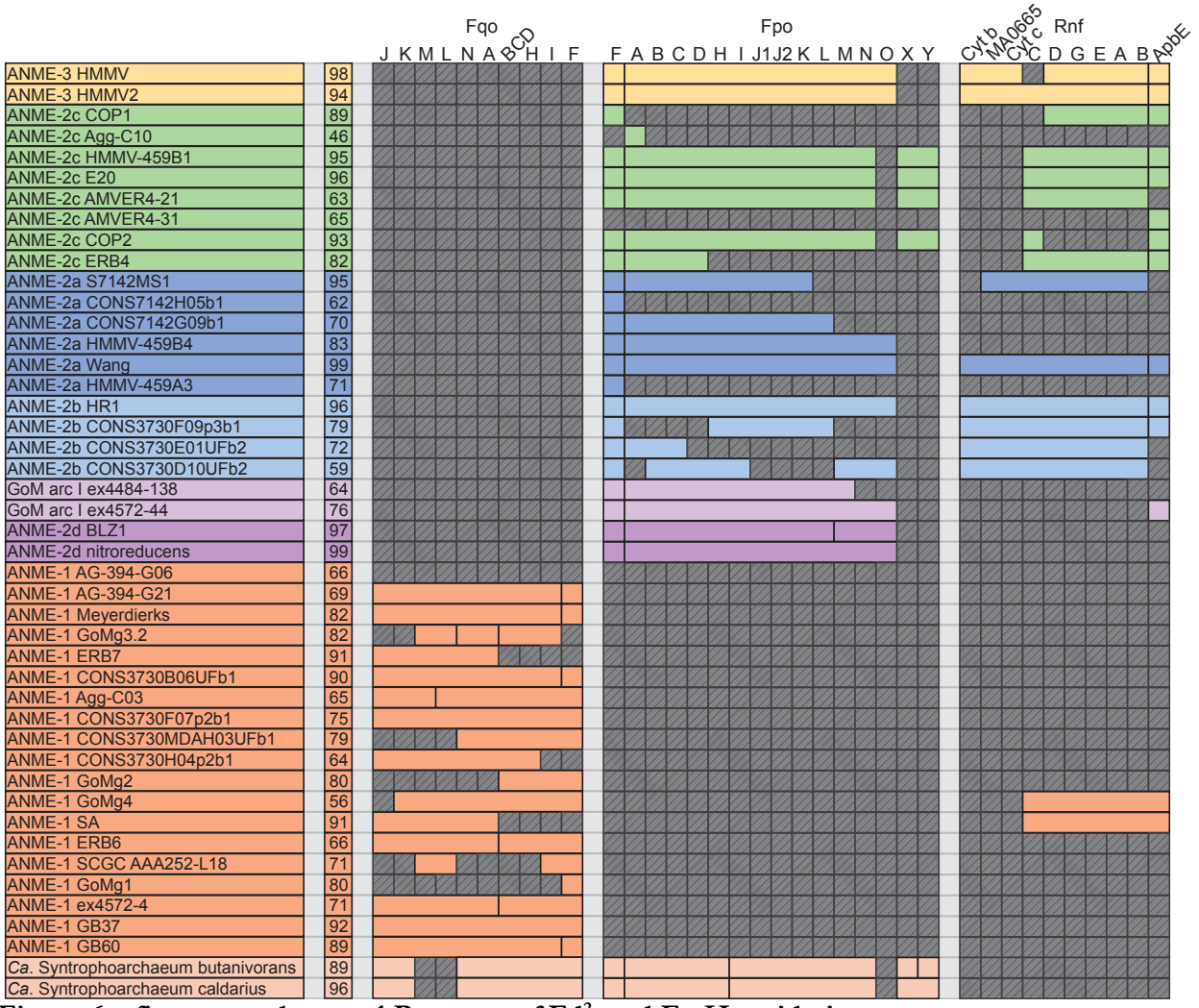


Figure 6—figure supplement 1 Presence of Fd^{2-} and $F_{420}H_2$ oxidation systems

The presence of each subunit of the oxidation complexes outlined in **Figure 6** is represented by colored boxes. ANME-2c and "*Ca.* Syntrophoarchaeum" lack the FpoO subunit found in other ANME and methanogens, and instead are followed by two highly expressed and conserved proteins denoted FpoX and Y. ApbE is a flavin transferase involved in Rnf biosynthesis. Missing genes are represented by transparent gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. Gene accession numbers can be found in **Figure 6—source data 1**.

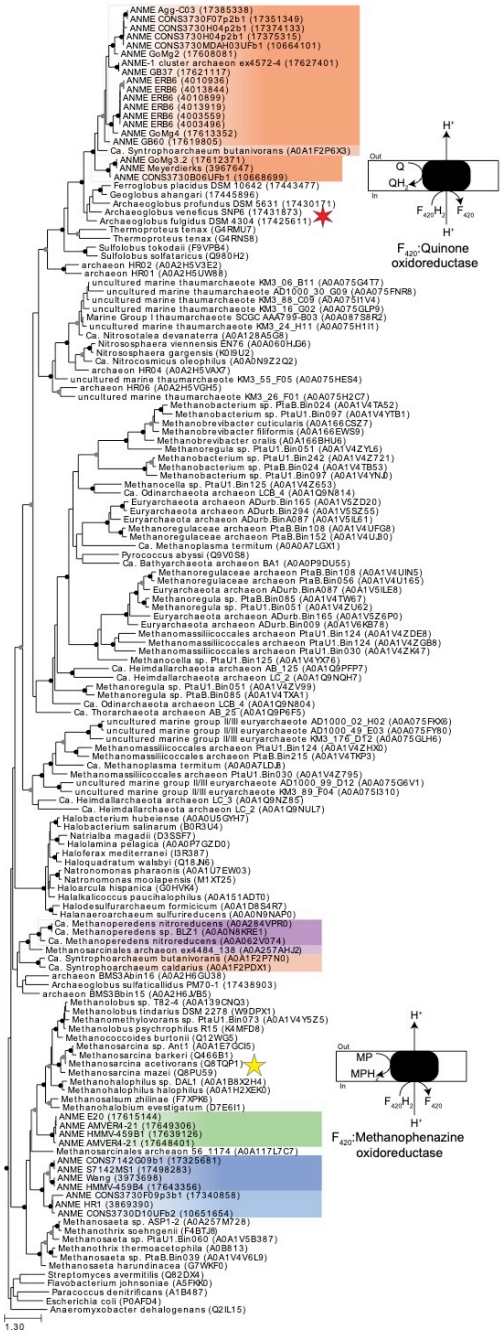


Figure 6—figure supplement 2 FpoH phylogenetic tree Phylogeny of the H subunit of Fpo/Fqo complexes found in ANME and related organisms demonstrates the large evolutionary distance between these complexes. See methods for details of tree construction. Sequences labeled with five and six-sided stars represent the biochemically characterized Fpo and Fqo complexes, respectively. ANME groups are highlighted. Closed circles represent branch support values of 80 to 100%, gray circles between 70 and 80%. Tree scales represent substitutions per site. Tree construction parameters are found in the methods section. Alignment and tree file can be found in **Figure 6—source data 1**.

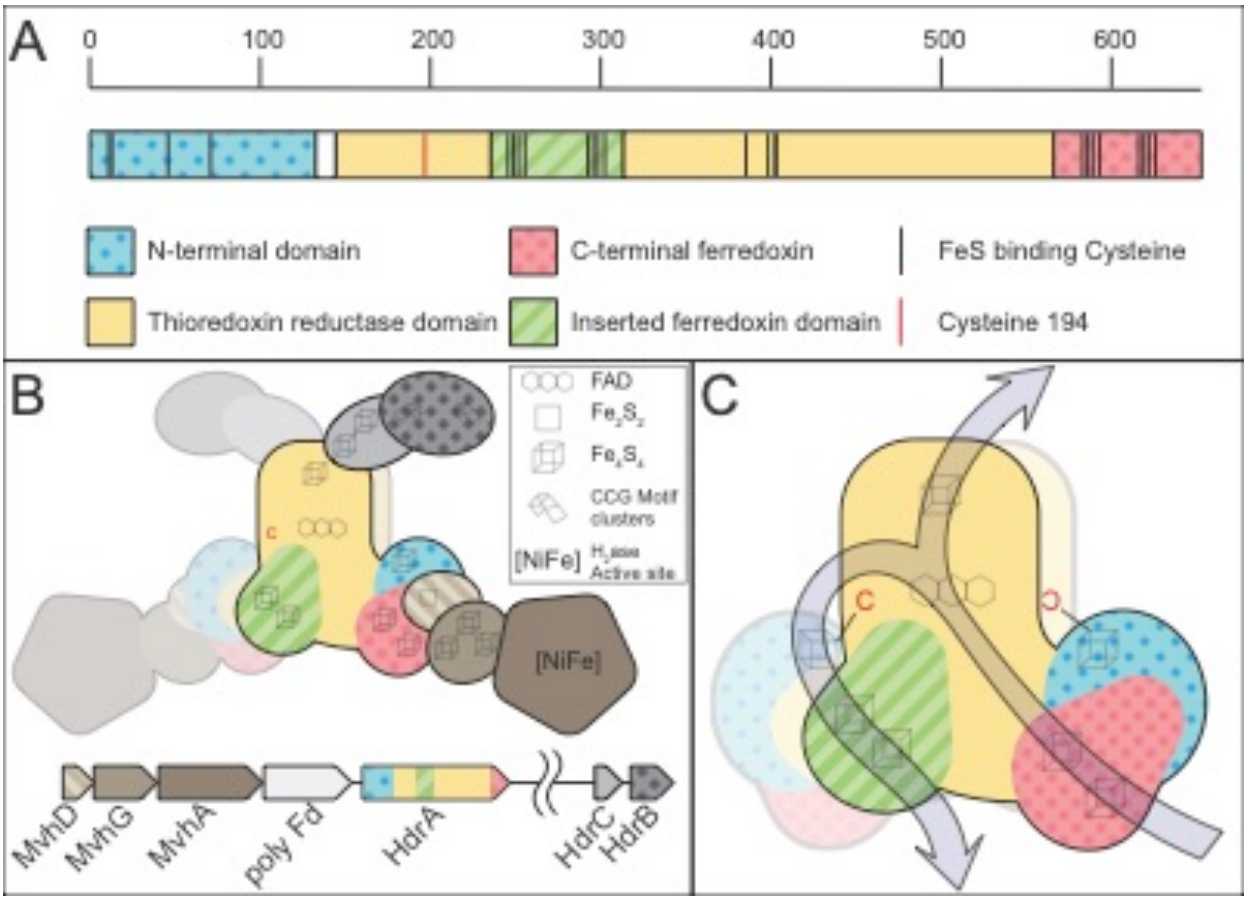


Figure 7 HdrABC structure overview Depiction of the primary structure of HdrA and the quarternary structure of the HdrABC-MvhADG complex based on the structure from *M. wolfeii.* (**A**) HdrA can be broken down into four domains, the positions of these domains and key iron-sulfur cluster binding cysteines are illustrated, scale denotes amino acid position in the *M. wolfeii* sequence. (**B**) Quarternary structure of the entire HdrABC-MvhABG complex illustrating the dimeric structure. (**C**) Detail of HdrA domain structure highlighting cofactor position and proposed electron flow from MvhD in through the C-terminal ferredoxin, bifurcation through the FAD cofactor, with two electrons flowing out through HdrBC via the Thioredoxin reductase domain's FeS cluster, while two other electrons flow out through the inserted ferredoxin domain, presumably to free ferredoxin (Fd^{2-}). Importantly for the proposed heterodimeric HdrA discussed here, this latter electron flow passes through the FeS cluster bound through a combination of Cys residues in the N-terminal domain, combined with a single Cys from the other HdrA subunit (Cys197 highlighted in red).

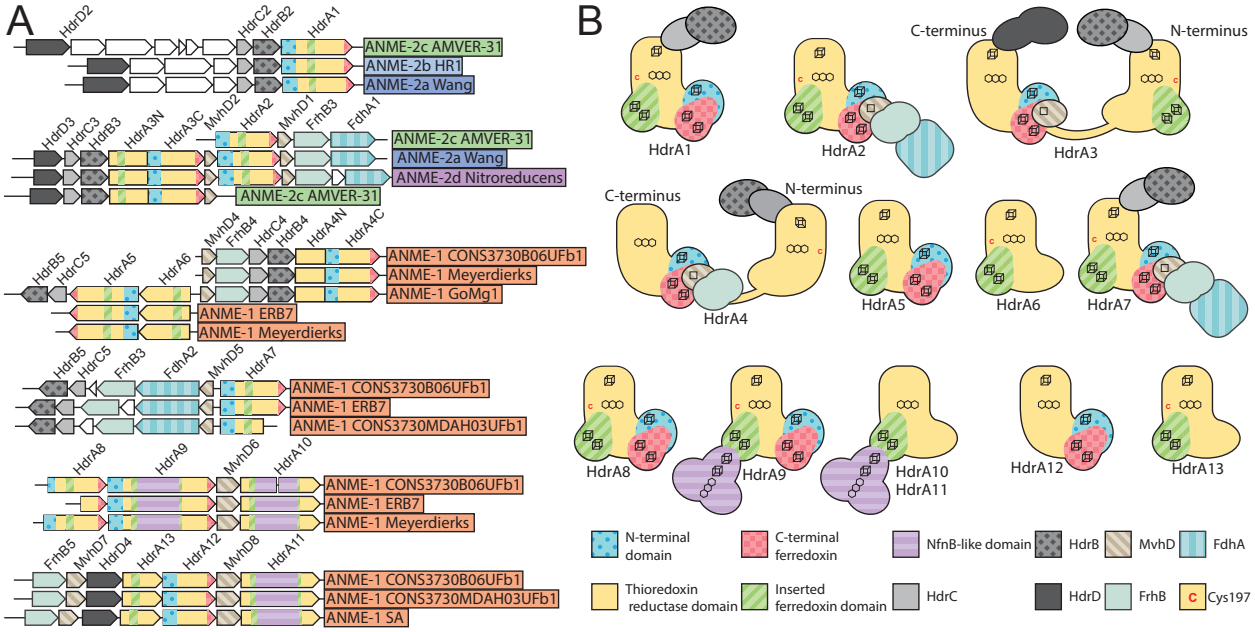


Figure 8 Hdr Operons and Domain heterogeneity (**A**) Examples of gene clusters containing HdrA genes from select ANME genomes. HdrA paralogs present in ANME have extensive modification to the domain structure as compared to the HdrA crystalized from *M. wolfeii* (see **Figure 7** for details of HdrA structure). These domains and associated protein subunits are illustrated with the gene context and orientation. (**B**) Illustration of conserved domains and cofactor binding residues in the 13 HdrA clusters defined here. All HdrA appeared to retain residues responsible for interaction with FAD, however the presence of FeS-binding cysteine residues, and entire domains as defined on the *M. wolfeii* structure are variably retained. Importantly, tandem or fused HdrA appear to have complementary presence/absence of C-terminal ferredoxin domains and Cys197 suggesting the formation of a heterodimeric complex.

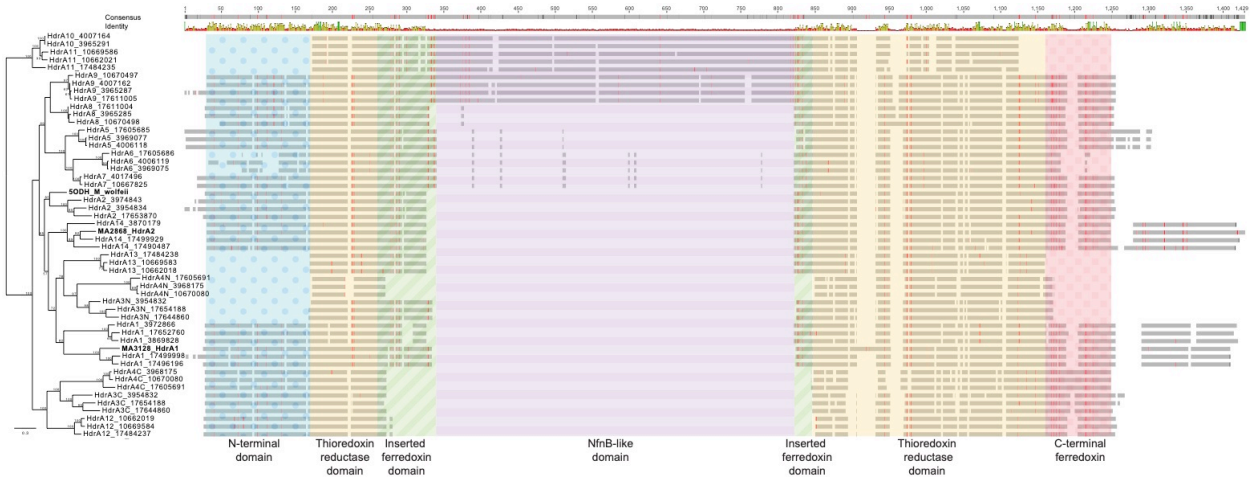


Figure 8—figure supplement 1 Alignment of HdrA sequences Protein alignments of representative ANME HdrA sequences were made and overlaid with the domains as colored in **Figure 7 and 8** based on the domains identified in the *M. wolfeii* structure. Cysteine residues are highlighted as red lines. Large insertion in the middle corresponds to the NfnB-like domain that interrupts the inserted ferredoxin domain in some ANME-1 HdrA. "HdrA1 (MA3128)" and "HdrA2 (MA2868)" from *M. acetivorans* (numbering from Buan and Metcalf, 2010) are included as references. Branch support values >50% are reported. Tree scales represent substitutions per site. Tree construction parameters are found in the methods section. Alignment and tree files can be found in **Figure 8—source data 1**.

Figure 8—figure supplement 2 Hdr counts per genome Most ANME and methanogen genomes contain 1-2 copies of *hdrABC* genes, however ANME-1 have a much greater abundance of *hdrA* homologs that are not accompanied by an increase in the number of *hdrB* or *C* homologs.

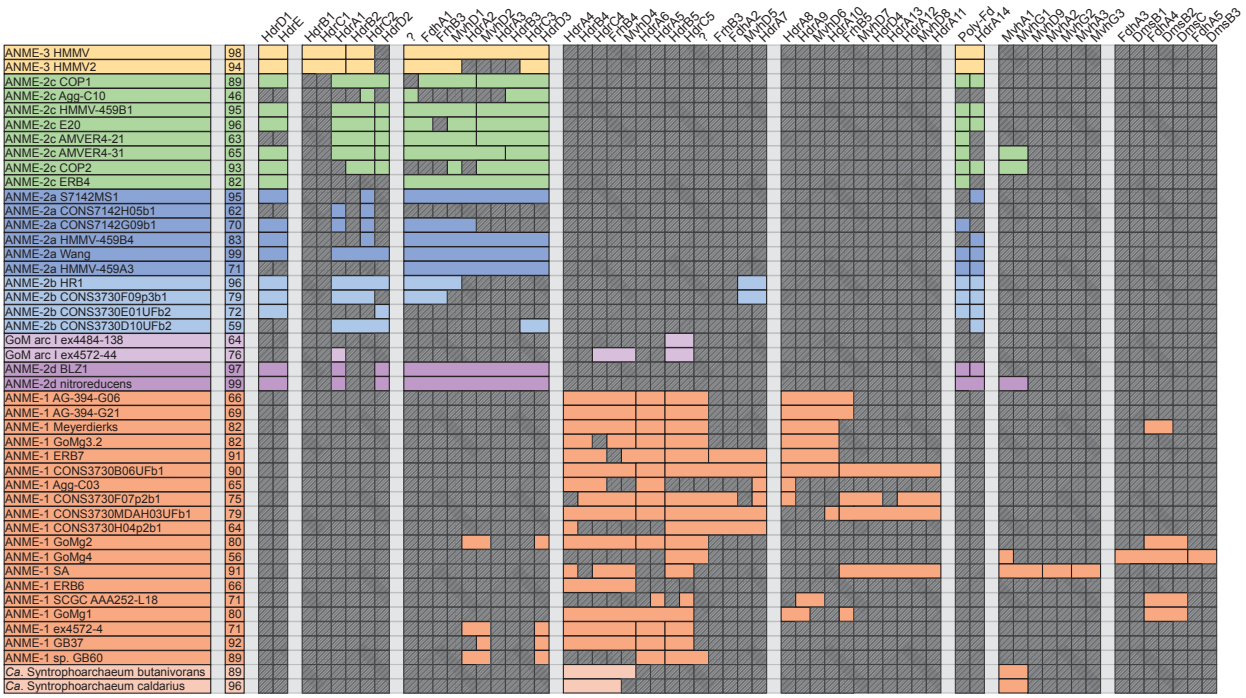


Figure 8—figure supplement 3 Hdr Presence/Absence Colored boxes represent presence of Hdr and associated genes in ANME genomes. HdrA1 in ANME is the same group as HdrA1 (MA3128) from *M. acetivorans*. Interestinly, only in ANME-3 is HdrA1 paired with HdrB1C1 (numbering from Buan and Metcalf, 2010). In other ANME HdrA1 is paired with HdrB2C2. HdrA2 from *M. acetivorans* (here called HdrA14), is found in ANME-2 and 3 and largely lacks clear gene clustering, although its associated polyferredoxin is found elsewhere in most genomes. Question marks represent genes with no known annotation but that are conserved in Hdr gene clusters. Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. Numbers were assigned to distinct clusters of paralogs, matching those labels found in **Figure 8**. Gene accession numbers can be found in **Figure 8—source data 1**.

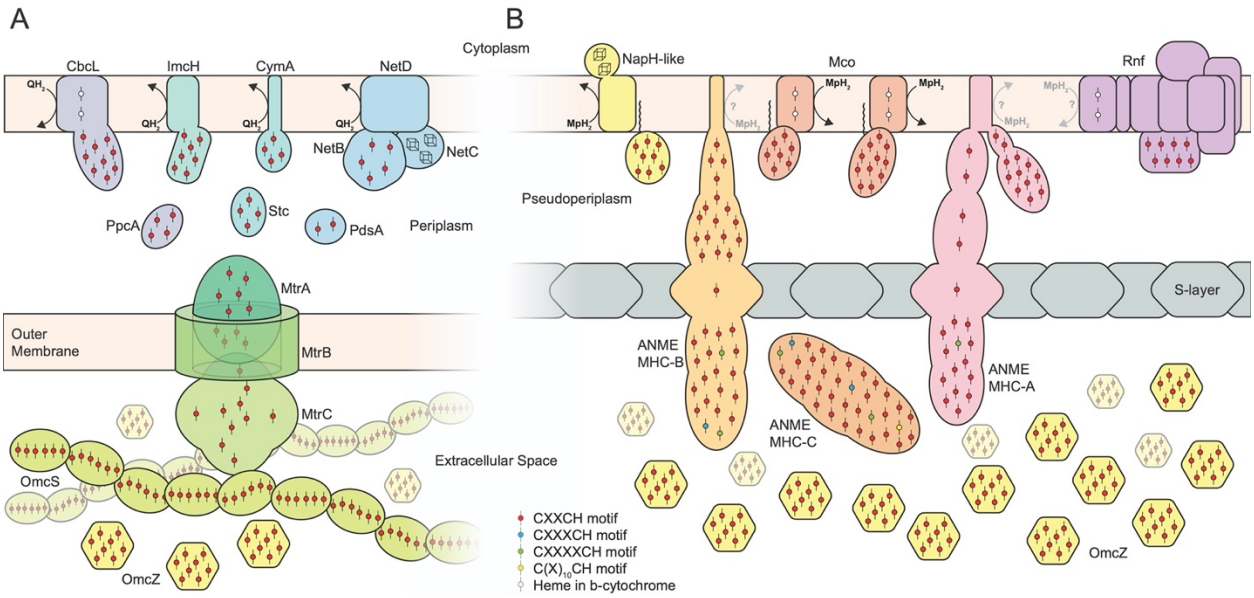


Figure 9 EET overview Comparison between EET systems known from gram negative bacteria and proposed analogous systems in ANME archaea. **A)** EET systems in gram negative bacteria involve membrane-bound quinol:cytochrome *c* oxidoreductases (CbcL, ImcH, CymA, NetD), small soluble cytochromes apparently involved in electron transport through the periplasmic space (PpcA, Stc, PdsA), and a beta-barrel/decaheme cytochrome *c* protein complex (MtrCAB) that acts as an electron conduit by which electrons can transit through the outer membrane to the extracellular space filled with additional cytochrome *c* such as OmcZ and filaments of OmcS. **B)** Analogous protein complexes found in ANME genomes that appear optimized for the challenges associated with EET in the archaeal cell architecture. MpH_2:cytochrome *c* oxidoreductases are likely present in the form of gene clusters containing VhtC cytochrome *b* subunits together with large 7 or 11 heme-binding MHC proteins (Mco). Other potential options for this step could include the NapH homologs sporadically distributed through ANME genomes, or through the action of the unique cytochrome *b* gene found in ANME Rnf clusters. Electron transfer through the outer proteinaceous S-layer requires a different mechanism than the beta-barrel/decaheme cytochrome strategy evolved in the EET-capable bacteria containing an outer membrane. This step is expected to be overcome by the giant ANME-specific MHC proteins containing S-layer domains allowing them to integrate into the S-layer structure. Very close homologs of OmcZ are found in ANME (see **Figure 10**). For details of S-layer MHC fusions, see **Figure 11**.

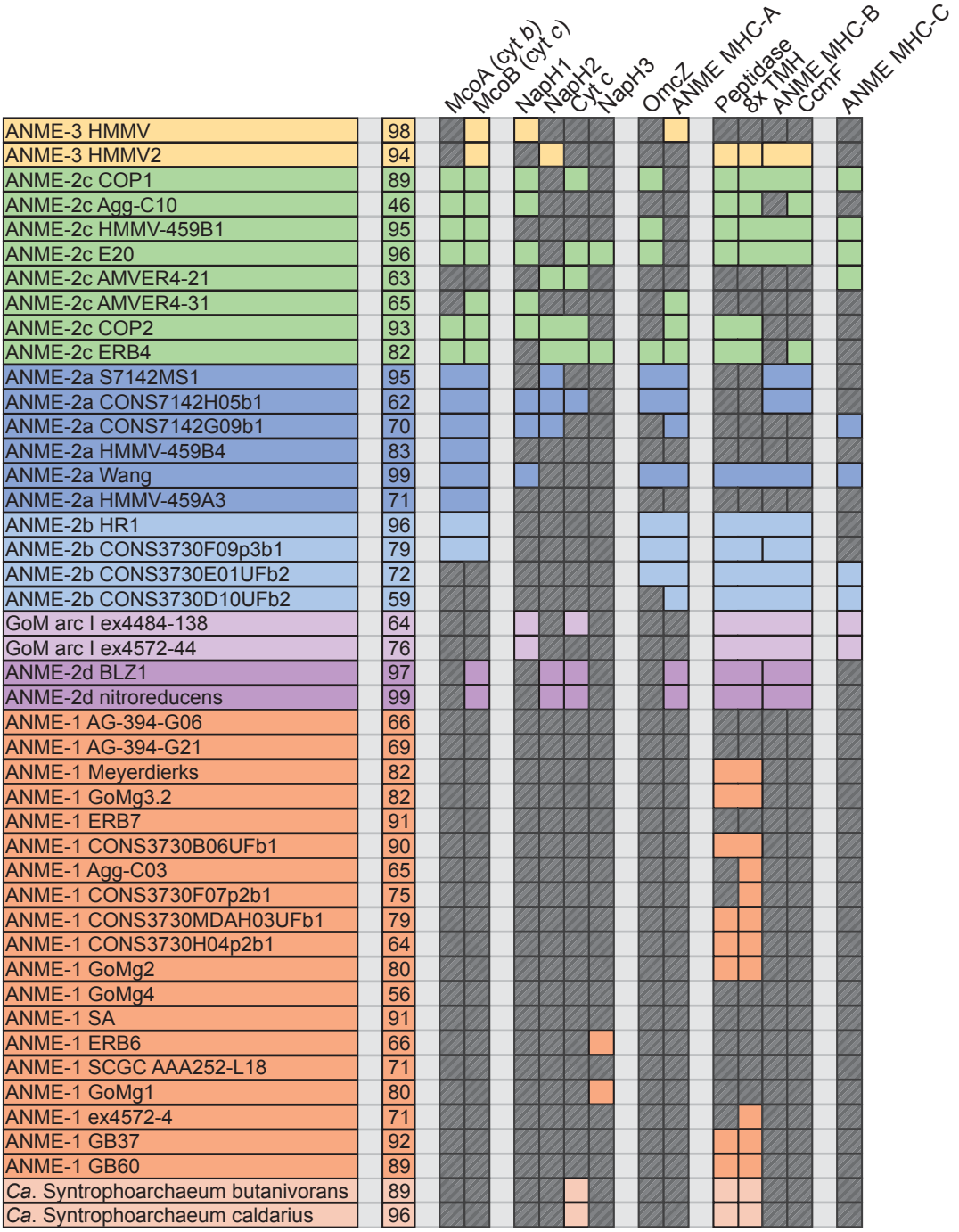


Figure 9—figure supplement 1 EET presence/absence Colored boxes represent presence of various key hypothetical EET genes in ANME genomes. Peptidase and 8x TMH denote a signal peptidase involved in cytochrome maturation and a hypothetical membrane protein with 8 predicted transmembrane helices that are found only in ANME and EET-capable *Ferroglobus* and *Geoglobus sp.* Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. Gene accession numbers can be found in **Figure 9—source data 1**.

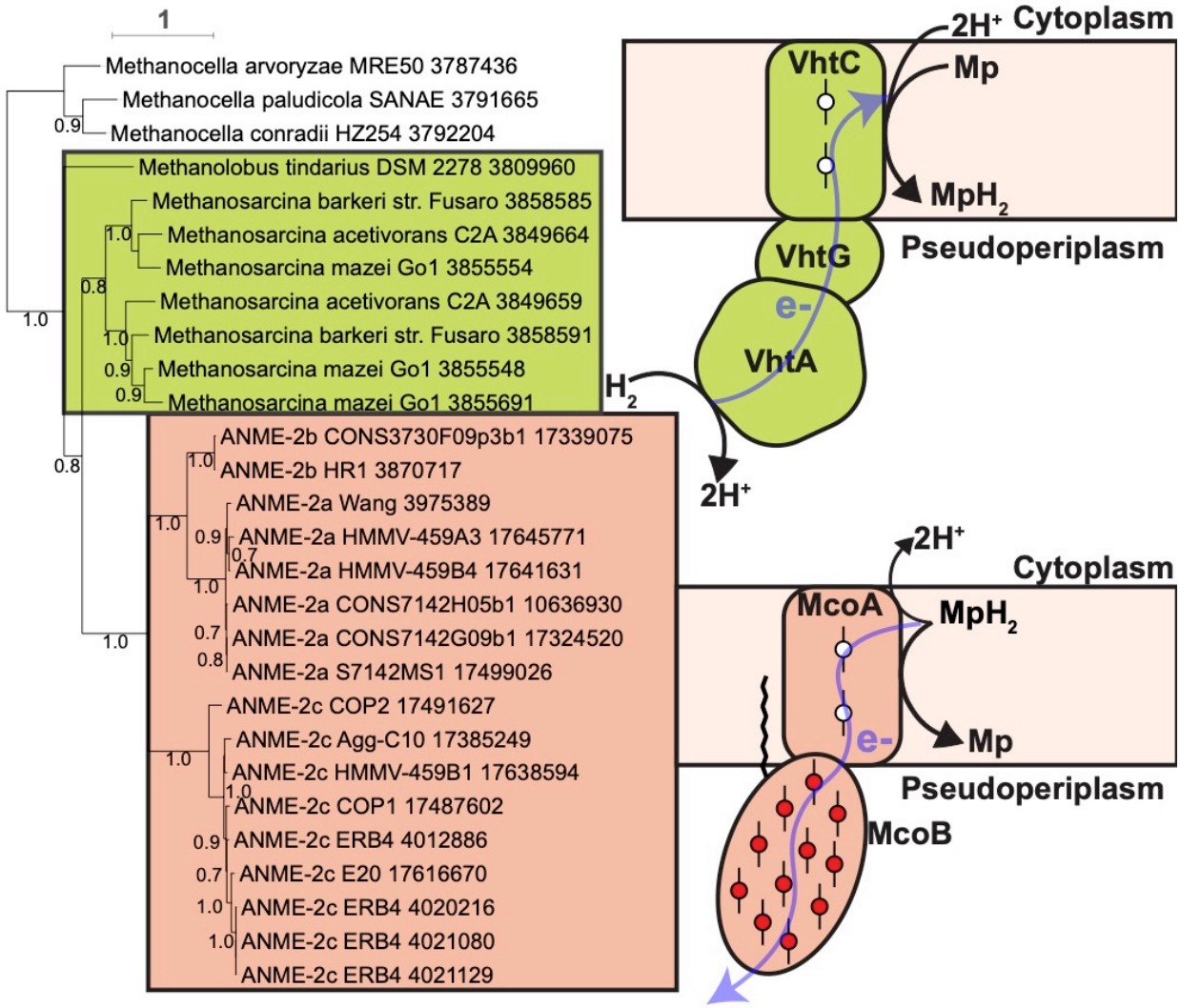


Figure 9—figure supplement 2 Mco phylogeny VhtC and McoA homologs from a selection of methanogens and the ANME genomes reported here were used to build a phylogenetic tree. ANME VhtC homologs clustered with multiheme cytochrome genes in ANME-2a and 2b, whereas in methanogens they always clustered with the VhtAG gene encoding the hydrogenase enzyme used by these organisms. Branch support values >50% are reported. Tree scales represent substitutions per site. Tree construction parameters are found in the methods section. Alignment and tree files can be found in **Figure 9—source data 1**.

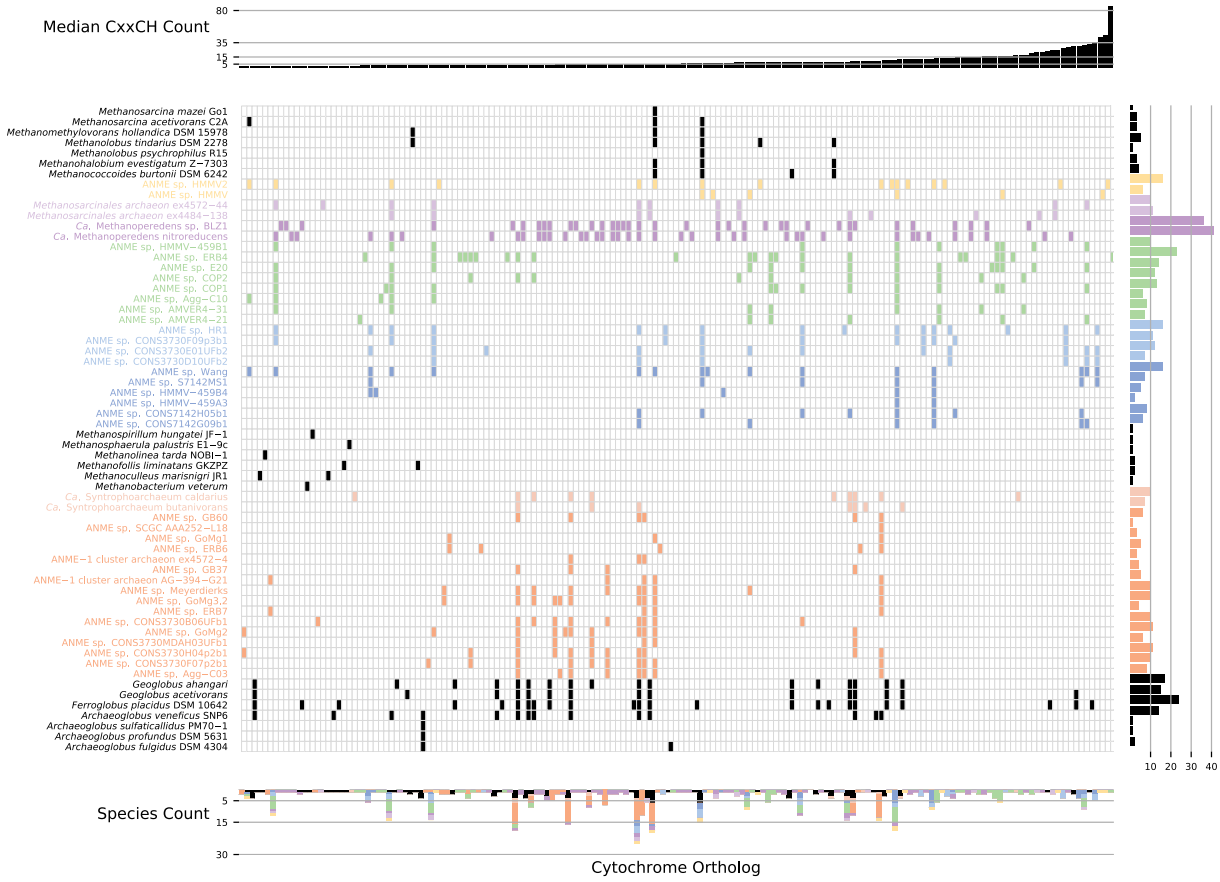


Figure 9—figure supplement 3 Cytochrome *c* orthologs Multiheme cytochromes present in ANME, methanogens and *Archeoglobales* genomes. The central panel depicts the presence of cytochromes between genomes, where cytochromes between genomes have been grouped based on reciprocal best-hit blast results. The cytochromes in the central panel are ordered based on the median number of heme binding sites (CxxCH motifs) present in each group (upper bar plot). The clade distribution of each cytochrome is shown in the lower bar plot, while the right bar plot depicts the number of cytochromes present in each organism. Genomes that did not contain any multiheme cytochromes, which were most methanogens, are not shown.

Figure 9—figure supplement 4 Broadly distributed cytochrome *c* homologs in ANME Unrooted phylogenetic reconstruction of the most broadly distributed group of MHC proteins found in the ANME archaea. Homologs to this group were only found in ANME, "*Ca.* Argoarchaeum", "*Ca.* Syntrophoarchaeum" and members of the *Archaeoglobales* known for their EET capabilities. Stars represents the sequence corresponding to the exceptionally highly expressed cytochromes highlighted in **Table 2**. Branch support values >90% are depicted as filled circles. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignment and tree files can be found in **Figure 9—source data 1**.

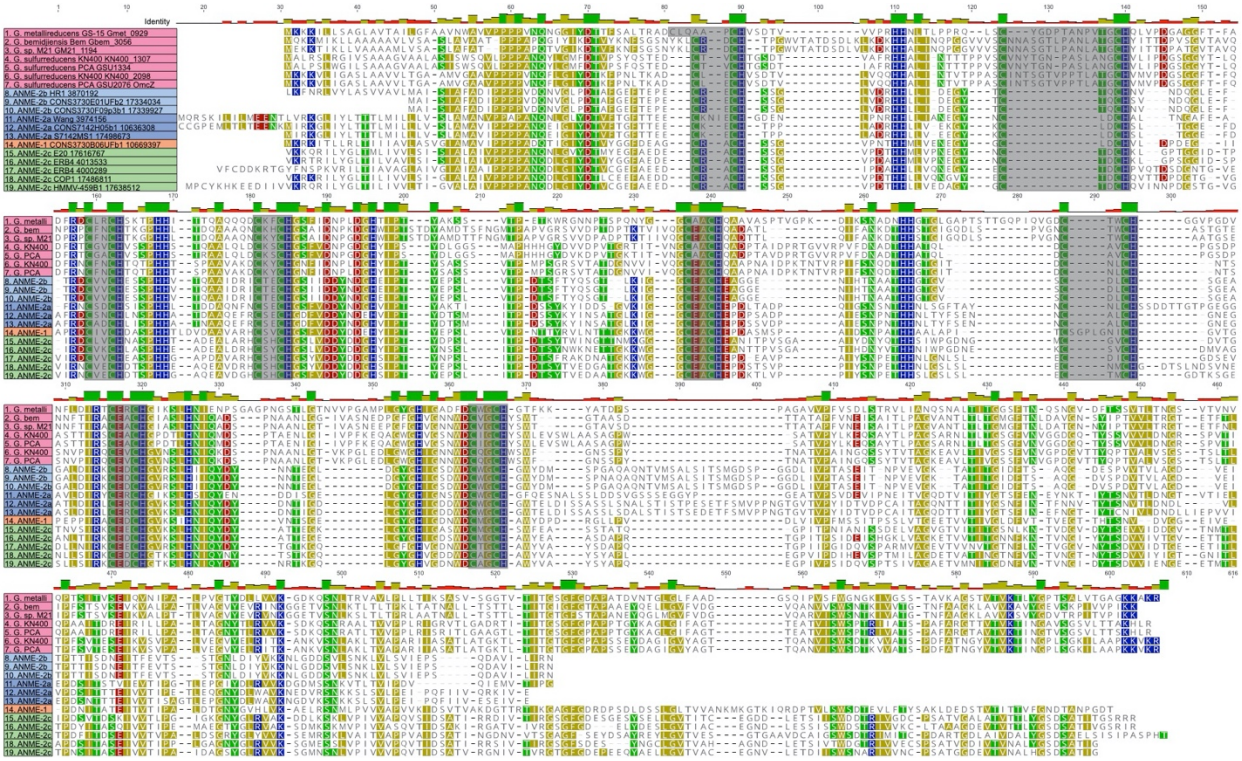


Figure 10 OmcZ homologs Protein sequence alignment of OmcZ homologs from various *Geobacter* species and ANME genomes reported here using muscle 3.8.31 with default settings. The eight CxxCH binding motifs are highlighted in gray. Regions of significant sequence identity are present throughout the protein, not just associated with the CxxCH motifs suggesting conserved function. Alignment file can be found in **Figure 10—source data 1**.

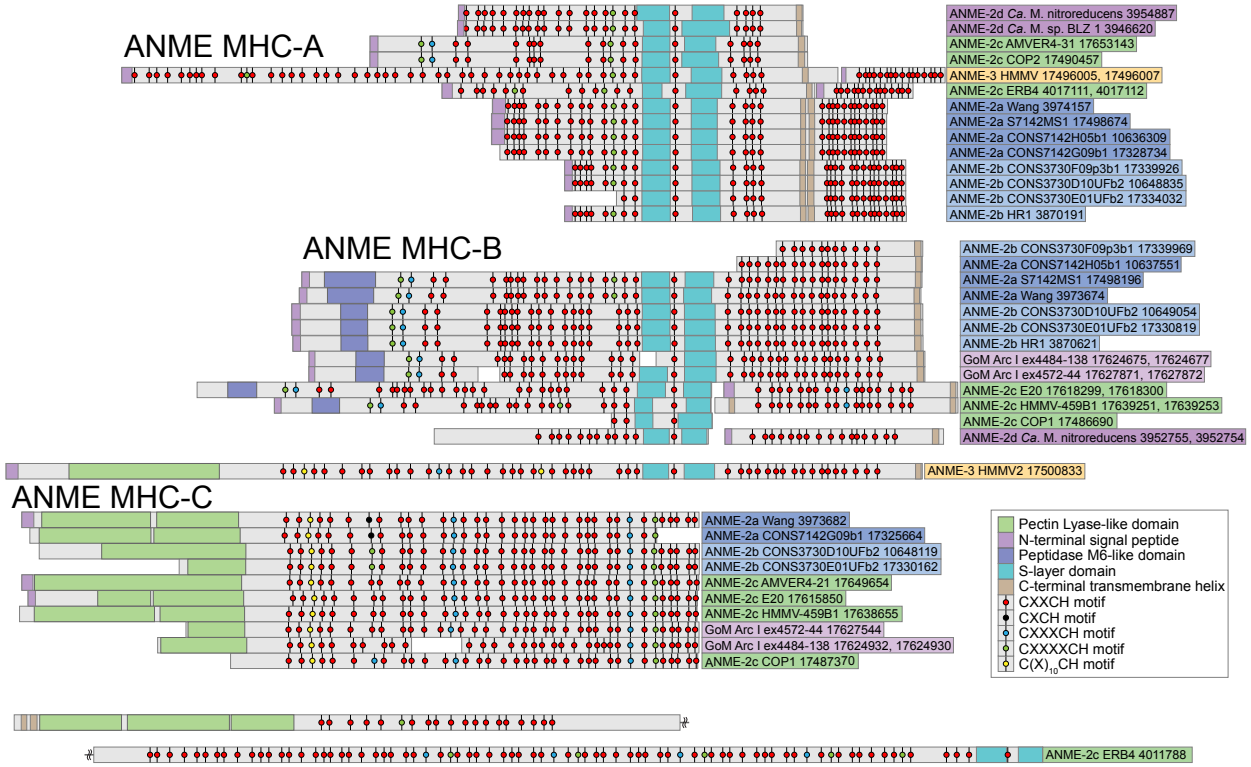


Figure 11 ANME MHC Schematic of protein structure highlighting the position of heme-binding motifs and other conserved features of the large ANME-specific multiheme cytochromes. The ANME MHC were divided into three major groups based on sequence similarity and conserved domains structure. ANME MHC-A contains an S-layer domain and C-terminal transmembrane helix. In ANME-2a and 2b these proteins are extended by an additional transmembrane helix and more heme-binding motifs. ANME MHC-B contains an S-layer domain and C-terminal transmembrane helix as well, but an N-terminal region devoid of heme-binding domains has similarity to peptidase M6-like domains. ANME MHC-C do not contain S-layers or C-terminal transmembrane helices, but instead contain a large N-terminal region with a predicted pectin lyase-type domain. Domains predicted with InterProScan and are displayed with colored boxes. Large MHC proteins from ANME-3 sp. HMMV2 and ANME-2c sp. ERB4 that do not clearly fit into these categories are also shown (Note: ANME-2c sp. ERB4 is a single peptide split between two lines due to its size).

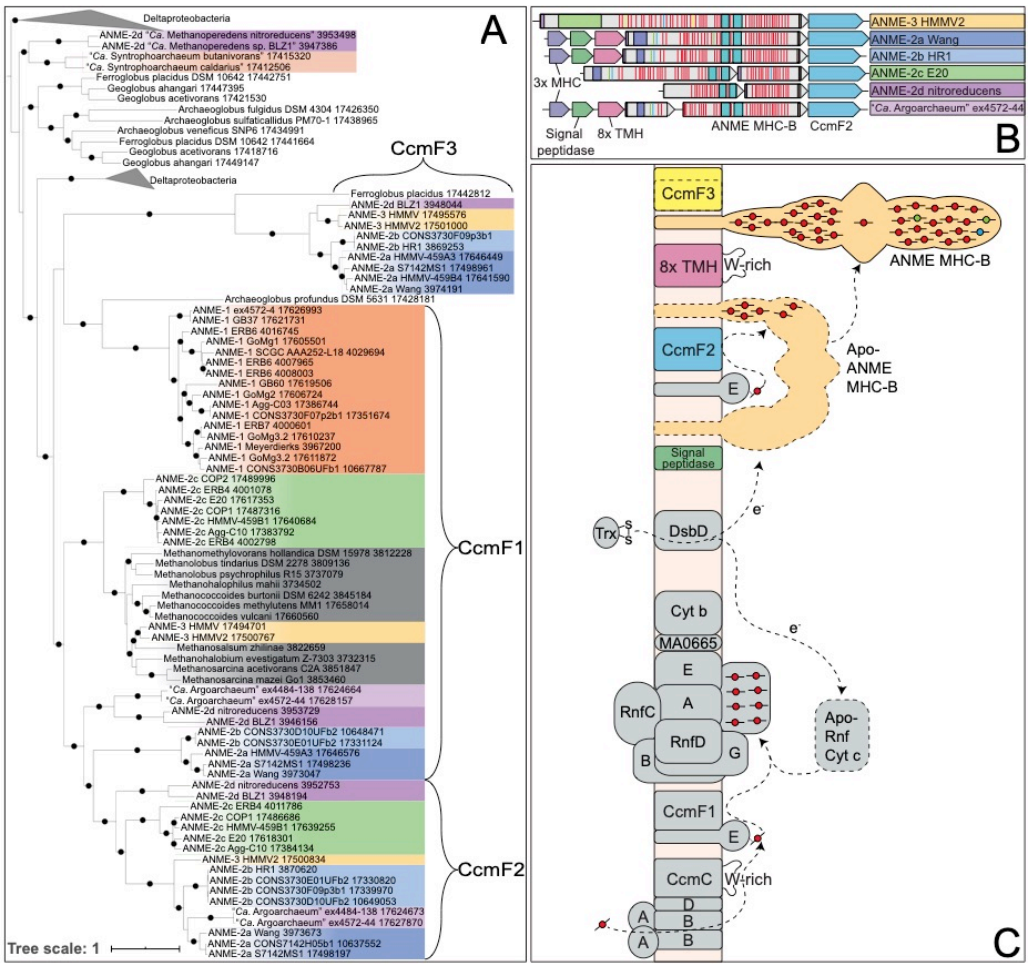


Figure 12 Cytochrome maturation and CcmF duplication (**A**) Phylogenetic analysis of CcmF homologs from ANME and closely related archaea. CcmF1 cluster contains the CcmF found in the *Methanosarcinaceae*. CcmF2 is a closely related group of homologs found only in ANME and "*Ca.* Argoarchaeum", and in all cases is found next to the S-layer containing ANME MHC-B. CcmF3 comprises the larger of two encoded genes that appear to be a split CcmF, and are found in ANME and *Ferroglobus*. (**B**) Example gene clusters from all groups that contain CcmF2 illustrating their position with respect to ANME MHC-B. Some genomes additionally contain a signal peptidase and a predicted membrane integral ANME-specific gene in this cluster (8x TMH). Horizontal red lines denote CxxCH heme binding domains, teal represent S-layer domain (see **Figure 11** for details of ANME-MHC structure). (**C**) Schematic of cytochrome maturation pathway. CcmA and B comprise ABC transporter module that exports heme B, which is transferred to CcmE via CcmC's tryptophan (W)-rich periplasmic loop. CcmE is expected to utilize CcmF1 to mature cytochrome *c* proteins found in both ANME and methanogens of the *Methanosarcinaceae*. CcmF2 found only next to ANME MHC-B is expected to be involved in its maturation. The co-occurring signal peptidase is likely involved in cleavage of the N-terminal signal sequence. Closed circles represent branch support values between 80 and 100%. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignment and tree files can be found in **Figure 12—source data 1**.

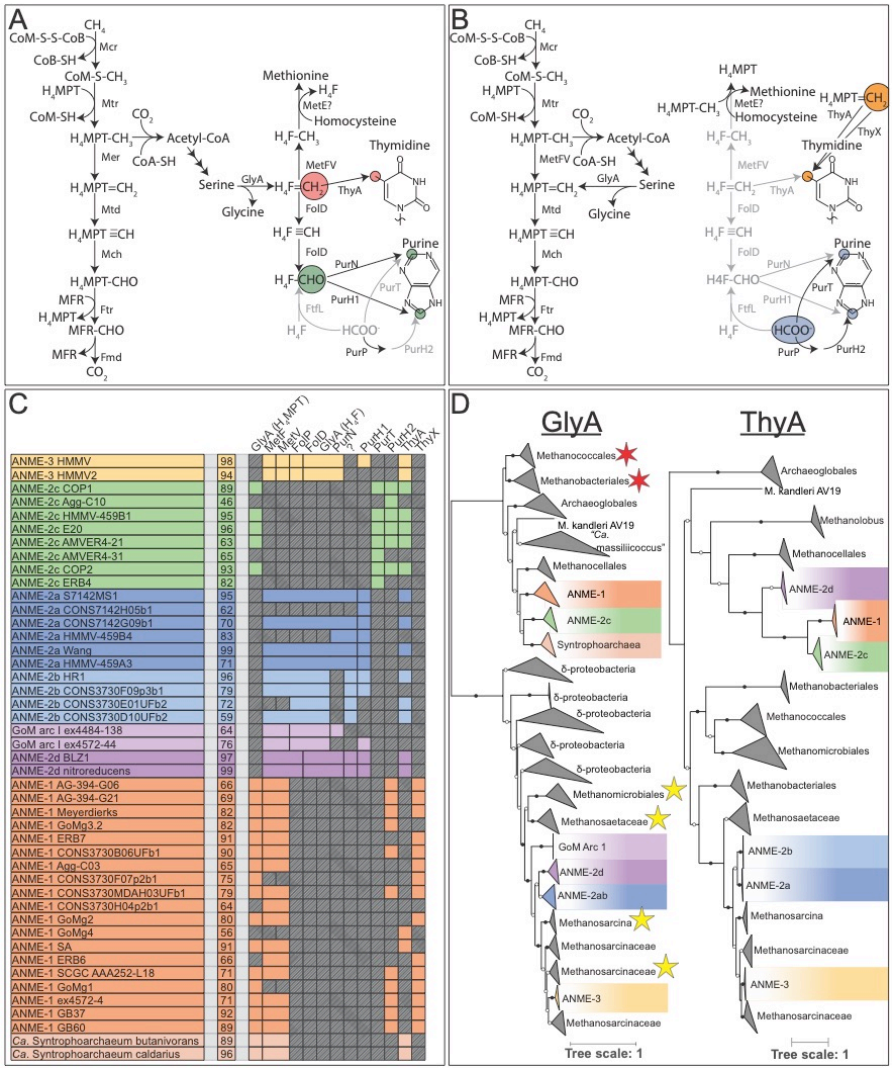


Figure 13 H_4MPT vs H_4F C1 metabolisms (**A**) Pathway of C1 transformations in ANME-2a, 2b, 2d and 3 based on presence of genes and precedent from *M. barkerii*. H_4MPT is the catabolic C1 carrier, while the H_4F C1 pool is derived from serine and is used for the biosynthesis of purine, thymidine and possibly methionine. (**B**) Pathway of C1 transformations available to ANME-1 and 2c which lack the vast majority of the H_4F interacting enzymes in (**A**). H_4F-CHO-specific enzymes in pyrimidine synthesis are replaced by free formate versions in ANME-2c. While present in these organisms, GlyA appears to be the type that interacts with H_4MPT instead of H_4F. (**C**) Colored boxes represent presence of various H_4F-interacting genes in ANME genomes. Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. (**D**) Phylogenetic analysis of GlyA and ThyA homologs found in ANME genomes. Red and yellow stars indicate GlyA sequences shown to react with H_4MPT and H_4F, respectively. Closed circles represent branch support values of 100%, open circles >50%. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignment and tree files can be found in **Figure 13—source data 1**.

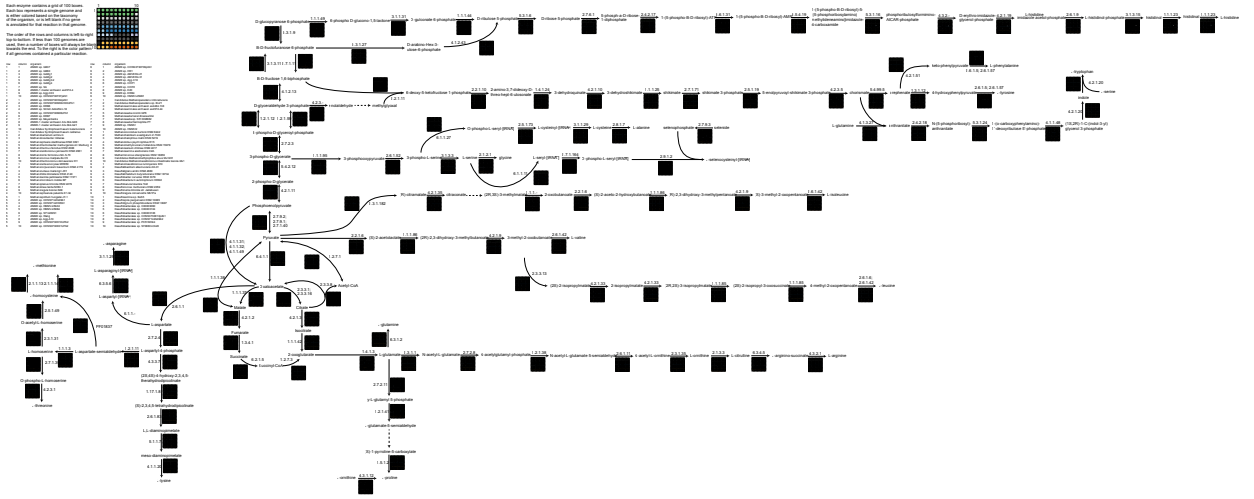


Figure 13—figure supplement 1 Amino acid pathways Amino acid pathways present in ANME genomes, methanogens and sulfate reducing bacteria. Each enzymatic reaction is labelled with an enzyme commission (EC) number, or in one case where an EC number is not available, a Pfam accession number and a matrix of small colored boxes. Each colored box represents a single genome, if the box is colored it means that that genome contains an annotation for that enzyme, otherwise it is left blank. Genomes are colored based on their clade.

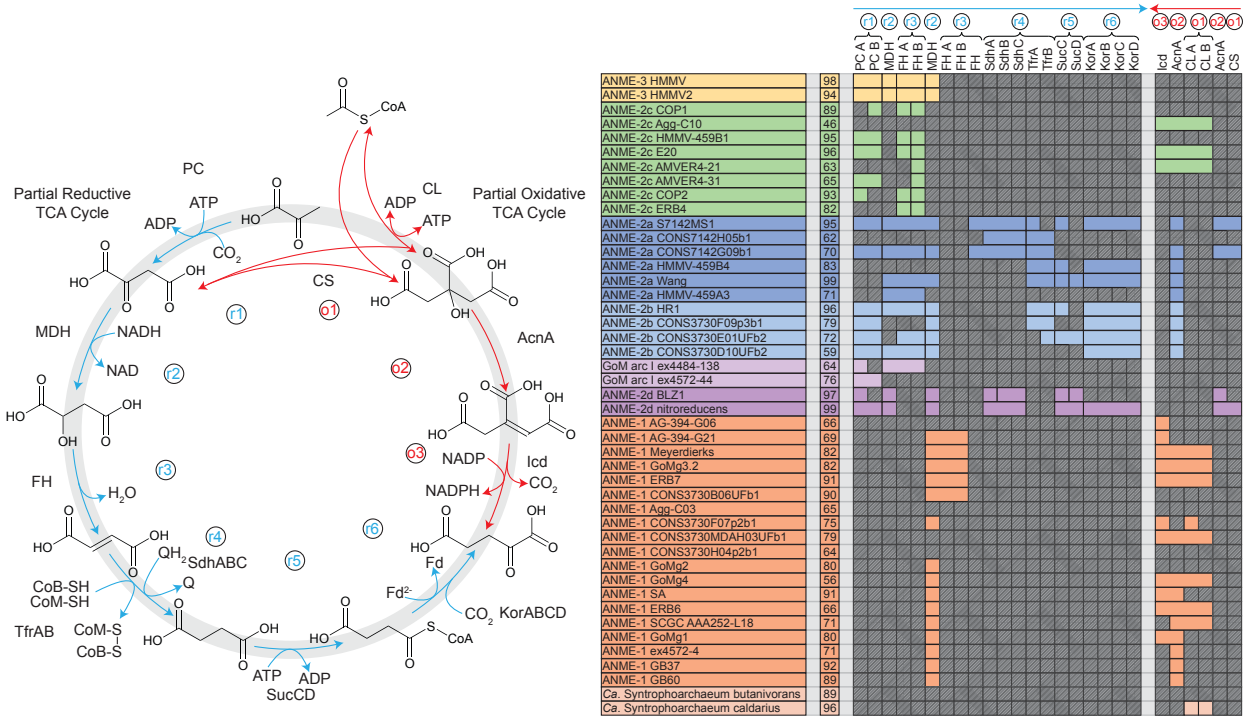


Figure 14 TCA cycle (**A**) Schematic of the enzymes and reactants in the partial oxidative and reductive TCA cycles in ANME that lead to 2-oxoglutarate for the purpose of producing anabolic intermediates for biosynthesis. Reductive pathway shown in blue, oxidative pathway shown in red. (**B**) Colored boxes represent presence of various TCA cycle genes in ANME genomes. Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. Note: some steps can be carried out by multiple different enzyme systems. Gene accession numbers can be found in **Figure 14—source data 1**.

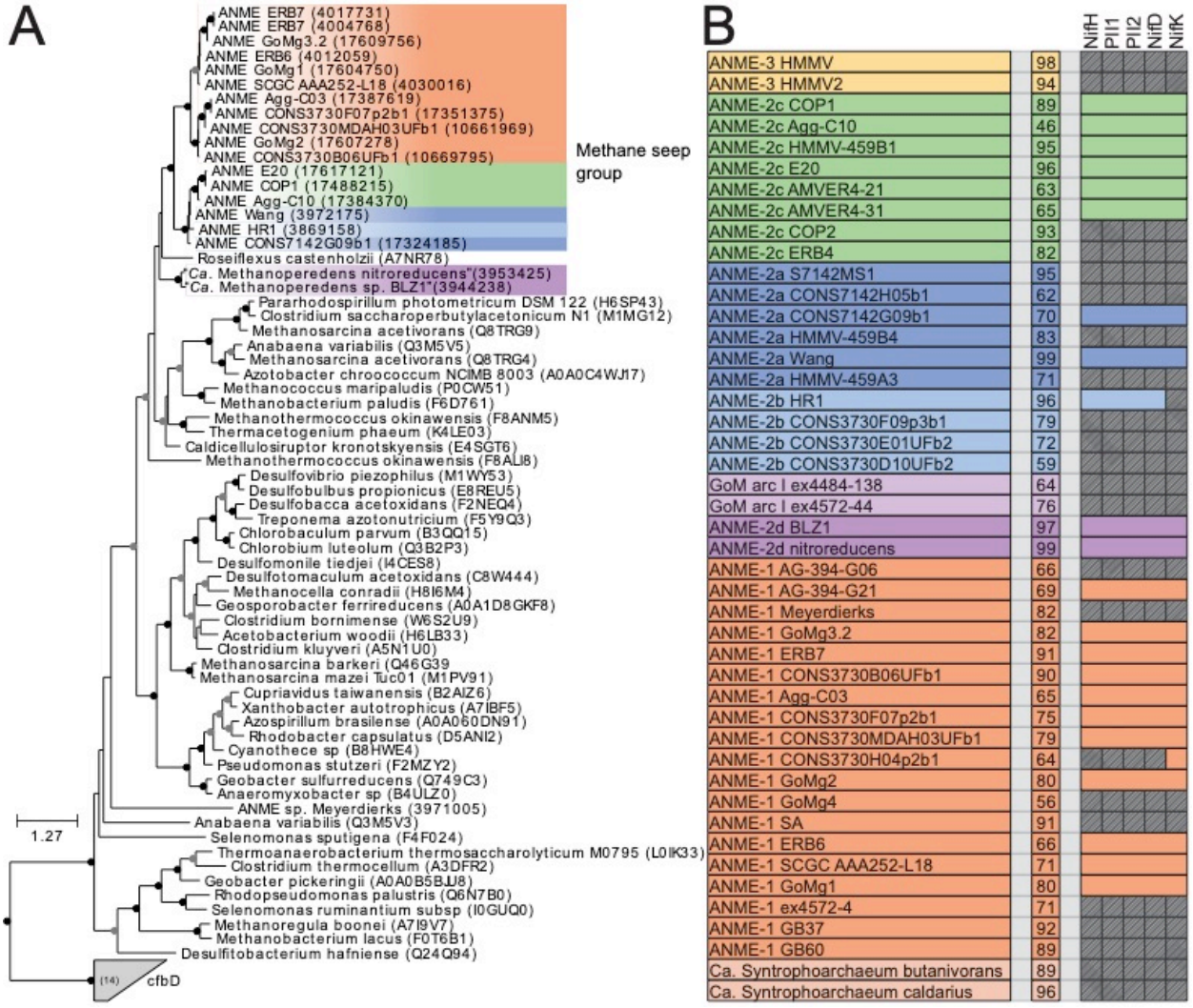


Figure 15 Methane seep group nitrogenase phylogeny and distribution (**A**) Maximum likelihood phylogenetic tree of NifD amino acid sequences from the "methane seep" group of nitrogenase, with close relatives. Closed circles represent branch support values of 80 to 100%, gray circles between 70 and 80%. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignments and tree files can be found in **Figure 15—source data 1**. (**B**) Presence of seep group nitrogenase in genomes presented here. Colored boxes represent presence of various nitrogenase related genes in ANME genomes. Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. Gene accession numbers can be found in **Figure 15—source data 1**.

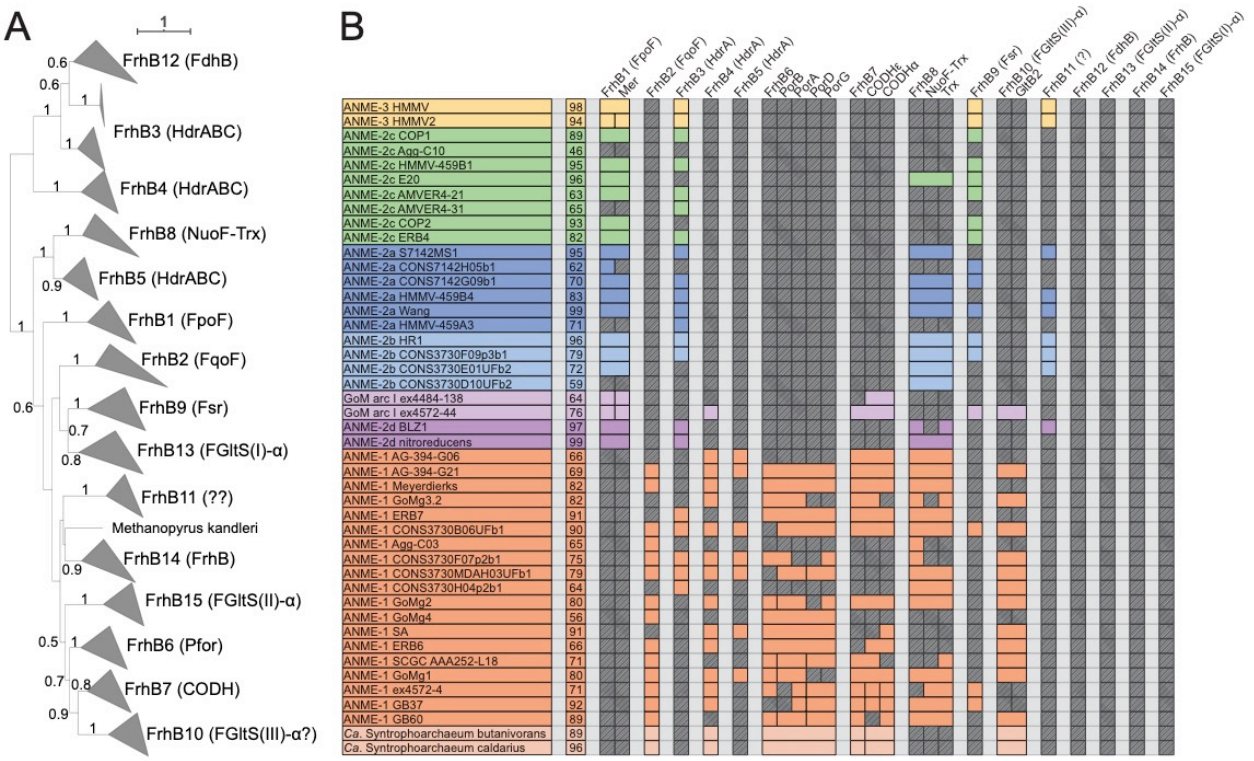


Figure 16 Diversity of FrhB family proteins Phylogeny and distribution of the FrhB paralogs recovered in our ANME genomes. (**A**) Phylogenetic tree built from all FrhB paralogs in ANME and methanogenic archaea. Major groups have been collapsed (for full tree see **Figure 16—figure supplement 1**). Label in parenthesis describe conserved gene cluster; FdhB: F_{420}-dependent formate dehydrogenase, FrhB: F_{420}-dependent hydrogenase, FqoF: F_{420}:quinone oxidoreductase, Fpo: F_{420}:phenazine oxidoreductase, HdrA: in cluster with HdrA genes, NuoF-Trx: in cluster with NuoF and thioredoxin genes, Fsr: F_{420}-dependent sulfite reductase, FGltS: F_{420}-dependent glutamate synthase, Pfor: in cluster with pyruvate ferredoxin oxidoreductase genes, CODH: in cluster with carbon monoxide dehydrogenase alpha and epsilon subunits. Only branch support values >50% are displayed. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignment and tree files can be found in **Figure 16—source data 1**. (**B**) Presence of various FrhB paralogs in different ANME genomes. Colored boxes represent presence of various nitrogenase related genes in ANME genomes. Missing genes are represented by gray boxes with diagonal line fill. Numbers in the second column represent genome completeness. When genes are together in a gene cluster their boxes are displayed fused together. Gene accession numbers can be found in **Figure 16—source data 1**.

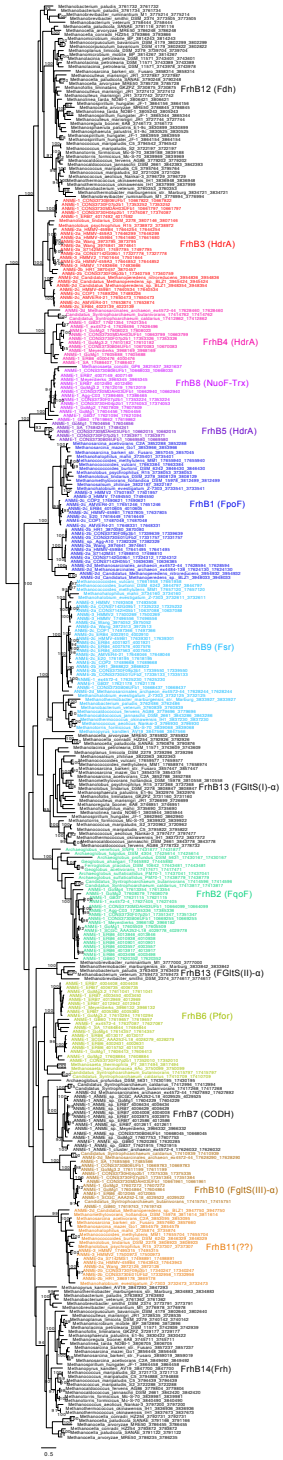


Figure 16—figure supplement 1 Expanded FdhB protein tree FrhB paralogs are labelled as in Figure 16. Only branch support values >90% are shown for clarity. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignment and tree files can be found in **Figure 16—source data 1**.

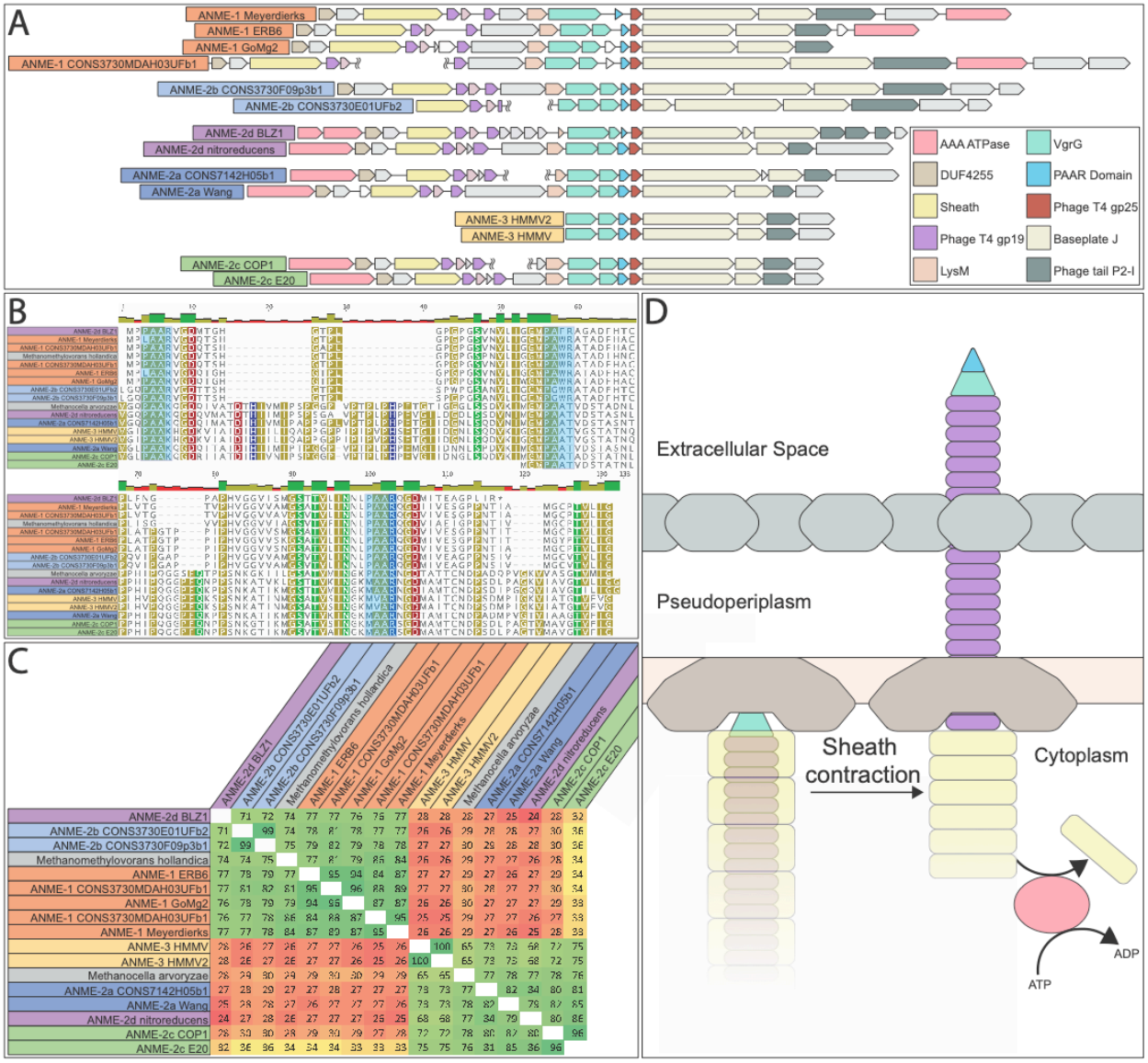


Figure 17 Phage-like protein translocation structures (**A**) PLTS gene clusters in ANME genomes. (**B**) Alignment of PAAR domain spike proteins. PAAR motifs highlighted in blue. Alignments were made using muscle 3.8.31 with default settings, and alignment file can be found in **Figure 17—source data 1** (**C**) amino acid identity of PAAR domain proteins highlighting two clear groupings. Note: closely related ANME-2a and 2b genomes contain PLTS structures belonging to different clusters. (**D**) Schematic of PLTS function.

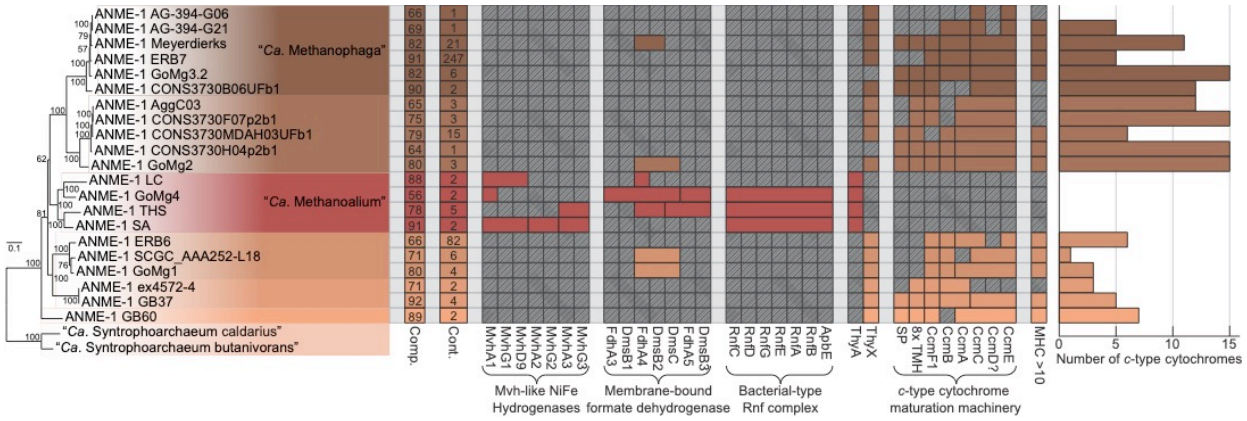


Figure 18 Comparison between "*Ca.* Methanoalium" and other ANME-1 genera Phylogenomic tree based on concatenated marker proteins highlighting individual ANME-1 genera, two of which have been assigned "*Candidatus*" names. Estimated genome completeness and contamination are shown in the first and second columns. Comparison of the presence of hydrogenase, membrane-associated formate dehydrogenases, Rnf complexes and cytochrome maturation machinery highlights important differences in electron flow between these genera. Genomes encoding MHC with more than 10 CxxCH heme-binding motifs are marked in the final column. Bar chart on the far right demonstrates the number of *c*-type cytochromes per genome. Only branch support values >50% are shown for clarity. Tree scales represent substitutions per site. Tree construction parameters are found in the Materials and Methods section. Alignment and tree files can be found in **Figure 18—source data 1**.

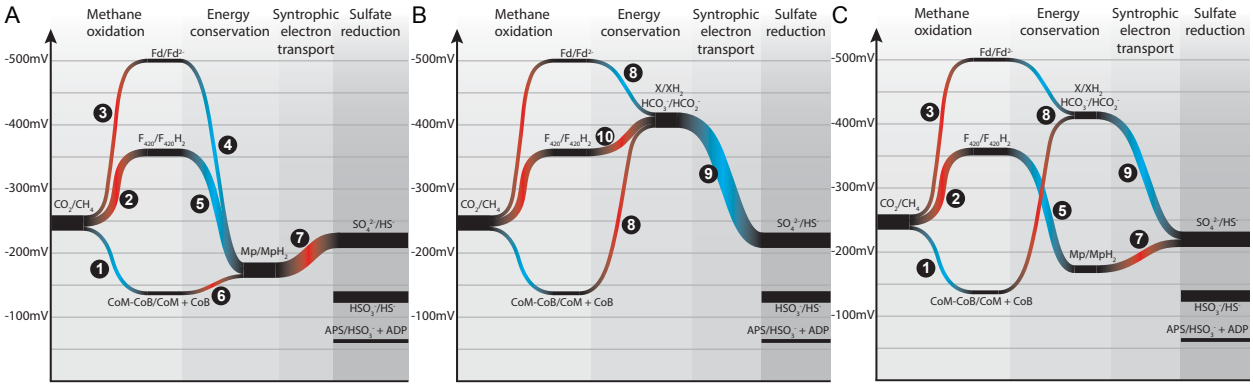


Figure 19 Energetics of mixed electron transfer Energy flow diagrams showing the redox potential (y-axis) of electrons as they travel from methane through the ANME energy metabolism to the SRB partner (x-axis is reaction progression). Width of paths correspond to number of electrons. Endergonic electron flow (uphill) shown in red, exergonic (downhill) in blue. Steps labelled with numbers are carried out by the following enzymes; 1: Mcr, 2: Mer/Mtd, 3: Fwd/Fmd, 4: Rnf, 5: Fpo/Fqo, 6: HdrDE, 7: Mco and ANME-MHC, 8: HdrABC and ANME-specific FrhB, FdhA, 9: uncatalyzed diffusion of electron carrier 10: FdhAB, possibly confurcation through HdrA complexes. X/XH_2 represent a hypothetical low-potential electron carrier that could be used as a diffusive electron shuttle. (**A**) electron transfer based entirely on EET from the methanophenazine pool. (**B**) electron transfer based entirely on soluble electron shuttle produced in the cytoplasmic space. (**C**) Mixed model with half of the electrons passing through each pathway.

References

Adam PS, Borrel G, Brochier-Armanet C, Gribaldo S. 2017. The growing tree of Archaea: new perspectives on their diversity, evolution and ecology. *ISME J*.

Albers S-V, Meyer BH. 2011. The archaeal cell envelope. *Nat Rev Microbiol* **9**:414–426.

Allen JWA, Harvat EM, Stevens JM, Ferguson SJ. 2006. A variant System I for cytochrome *c* biogenesis in archaea and some bacteria has a novel CcmE and no CcmH. *FEBS Lett* **580**:4827–4834. doi:10.1016/j.febslet.2006.07.073

Allen KD, Miller DV, Rauch BJ, Perona JJ, White RH. 2015. Homocysteine is biosynthesized from aspartate semialdehyde and hydrogen sulfide in methanogenic archaea. *Biochemistry* **54**:3129–3132.

Angelaccio S, Chiaraluce R, Consalvi V, Buchenau B, Giangiacomo L, Bossa F, Contestabile R. 2003. Catalytic and Thermodynamic Properties of Tetrahydromethanopterin-dependent Serine Hydroxymethyltransferase from *Methanococcus jannaschii*. *J Biol Chem* **278**:41789–41797. doi:10.1074/jbc.M306747200

Anthony C. 1982. The biochemistry of methylotrophs. London: Academic press.

Arbing MA, Chan S, Shin A, Phan T, Ahn CJ, Rohlin L, Gunsalus RP. 2012. Structure of the surface layer of the methanogenic archaean Methanosarcina acetivorans. *Proc Natl Acad Sci* **109**:11812–11817. doi:10.1073/pnas.1120595109

Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S,

Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* **36**:566–569. doi:10.1038/nbt.4163

Arshad A, Speth DR, de Graaf RM, Op den Camp HJM, Jetten MSM, Welte CU. 2015. A Metagenomics-Based Metabolic Model of Nitrate-Dependent Anaerobic Oxidation of Methane by Methanoperedens-Like Archaea. *Front Microbiol* **6**:1423.

Basta T, Boum Y, Briffotaux J, Becker HF, Lamarre-Jouenne I, Lambry J-C, Skouloubris S, Liebl U, Graille M, van Tilbeurgh H, Myllykallio H. 2012. Mechanistic and structural basis for inhibition of thymidylate synthase ThyX. *Open Biol* **2**:120120–120120. doi:10.1098/rsob.120120

Baumer S, Ide T, Jacobi C, Johann A, Gottschalk G, Deppenmeier U. 2000. The F420H2 dehydrogenase from Methanosarcina mazei is a Redox-driven proton pump closely related to NADH dehydrogenases. *J Biol Chem* **275**:17968–17973.

Bayer EA, Belaich J-P, Shoham Y, Lamed R. 2004. The Cellulosomes: Multienzyme Machines for Degradation of Plant Cell Wall Polysaccharides. *Annu Rev Microbiol* **58**:521–554. doi:10.1146/annurev.micro.57.030502.091022

Bayer EA, Coutinho PM, Henrissat B. 1999. Cellulosome-like sequences in *Archaeoglobus fulgidus* : an enigmatic vestige of cohesin and dockerin domains. *FEBS Lett* **463**:277–280. doi:10.1016/S0014-5793(99)01634-8

Beal EJ, House CH, Orphan VJ. 2009. Manganese- and Iron-Dependent Marine Methane Oxidation. *Science* **325**:184–187. doi:10.1126/science.1169984

Berg IA, Kockelkorn D, Ramos-Vera WH, Say RF, Zarzycki J, Hügler M, Alber BE, Fuchs G. 2010. Autotrophic carbon fixation in archaea. *Nat Rev Microbiol* **8**:447–460. doi:10.1038/nrmicro2365

Berk H, Thauer RK. 1998. F420H2:NADP oxidoreductase from Methanobacterium thermoautotrophicum: identification of the encoding gene via functional overexpression in Escherichia coli. *FEBS Lett* **438**:124–126. doi:10.1016/S0014-5793(98)01288-5

Berk H, Thauer RK. 1997. Function of coenzyme F 420 -dependent NADP reductase in methanogenic archaea containing an NADP-dependent alcohol dehydrogenase. *Arch Microbiol* **168**:396–402. doi:10.1007/s002030050514

Berntsson RP-A, Doeven MK, Fusetti F, Duurkens RH, Sengupta D, Marrink S-J, Thunnissen A-MWH, Poolman B, Slotboom D-J. 2009. The structural basis for peptide selection by the transport receptor OppA. *EMBO J* **28**:1332–1340. doi:10.1038/emboj.2009.65

Bertram S, Blumenberg M, Michaelis W, Siegert M, Krüger M, Seifert R. 2013. Methanogenic capabilities of ANME-archaea deduced from ^{13}C-labelling approaches. *Environ Microbiol* **15**:2384–2393. doi:10.1111/1462-2920.12112

Beulig F, Røy H, McGlynn SE, Jørgensen BB. 2019. Cryptic CH4 cycling in the sulfate–methane transition of marine sediments apparently mediated by ANME-1 archaea. *ISME J* **13**:250–262. doi:10.1038/s41396-018-0273-z

Beveridge TJ, Sprott GD, Whippey P. 1991. Ultrastructure, inferred porosity, and gram-staining character of Methanospirillum hungatei filament termini describe a unique cell permeability for this archaeobacterium. *J Bacteriol* **173**:130–140. doi:10.1128/jb.173.1.130-140.1991

Biegel E, Schmidt S, González JM, Müller V. 2011. Biochemistry, evolution and physiological function of the Rnf complex, a novel ion-motive electron transport complex in prokaryotes. *Cell Mol Life Sci* **68**:613–634.

Bock A-K, Kunow J, Glasemacher J, Schönheit P. 1996. Catalytic Properties, Molecular Composition and Sequence Alignments of Pyruvate: Ferredoxin Oxidoreductase from the Methanogenic Archaeon Methanosarcina Barkeri (Strain Fusaro). *Eur J Biochem* **237**:35–44. doi:10.1111/j.1432-1033.1996.0035n.x

Boetius A, Ravenschlag K, Schubert CJ, Rickert D, Widdel F, Gieseke A, Amann R, Jørgensen BB, Witte U, Pfannkuche O. 2000. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature* **407**:623–626. doi:10.1038/35036572

Borrel G, Adam PS, McKay LJ, Chen L-X, Sierra-García IN, Sieber CMK, Letourneur Q, Ghozlane A, Andersen GL, Li W-J, Hallam SJ, Muyzer G, Oliveira VM de, Inskeep WP, Banfield JF, Gribaldo S. 2019. Wide diversity of methane and short-chain alkane metabolisms in uncultured archaea. *Nat Microbiol* **4**:603–613. doi:10.1038/s41564-019-0363-3

Bowles M, Joye S. 2011. High rates of denitrification and nitrate removal in cold seep sediments. *ISME J* **5**:565–567. doi:10.1038/ismej.2010.134

Boyd JA, Jungbluth SP, Leu AO, Evans PN, Woodcroft BJ, Chadwick GL, Orphan VJ, Amend JP, Rappé MS, Tyson GW. 2019. Divergent methyl-coenzyme M reductase genes in a deep-subseafloor Archaeoglobi. *ISME J* **13**:1269–1279. doi:10.1038/s41396-018-0343-2

Brazelton WJ, Ludwig KA, Sogin ML, Andreishcheva EN, Kelley DS, Shen C-C, Edwards RL, Baross JA. 2010. Archaea and bacteria with surprising microdiversity show shifts in dominance over 1,000-year time scales in hydrothermal chimneys. *Proc Natl Acad Sci* **107**:1612–1617. doi:10.1073/pnas.0905369107

Brazelton WJ, Schrenk MO, Kelley DS, Baross JA. 2006. Methane- and Sulfur-Metabolizing Microbial Communities Dominate the Lost City Hydrothermal Field Ecosystem. *Appl Env Microbiol* **72**:6257–6270. doi:10.1128/AEM.00574-06

Brown AM, Hoopes SL, White RH, Sarisky CA. 2011. Purine biosynthesis in archaea: variations on a theme. *Biol Direct* **6**:63.

Brüggemann H, Falinski F, Deppenmeier U. 2000. Structure of the F420H2:quinone oxidoreductase of Archaeoglobus fulgidus identification and overproduction of the F420H2-oxidizing subunit. *Eur J Biochem* **267**:5810–5814.

Brutinel ED, Gralnick JA. 2012. Shuttling happens: soluble flavin mediators of extracellular electron transfer in Shewanella. *Appl Microbiol Biotechnol* **93**:41–48. doi:10.1007/s00253-011-3653-0

Buan NR, Metcalf WW. 2010. Methanogenesis by Methanosarcina acetivorans involves two structurally and functionally distinct classes of heterodisulfide reductase. *Mol Microbiol* **75**:843–853. doi:10.1111/j.1365-2958.2009.06990.x

Buchenau B, Thauer RK. 2004. Tetrahydrofolate-specific enzymes in Methanosarcina barkeri and growth dependence of this methanogenic archaeon on folic acid or p-aminobenzoic acid. *Arch Microbiol* **182**:313–325.

Buchfink B, Xie C, Huson DH. 2015. Fast and sensitive protein alignment using DIAMOND. *Nat Methods* **12**:59–60.

Buckel W, Thauer RK. 2018. Flavin-Based Electron Bifurcation, A New Mechanism of Biological Energy Coupling. *Chem Rev* **118**:3862–3886. doi:10.1021/acs.chemrev.7b00707

Buckel W, Thauer RK. 2013. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na+ translocating ferredoxin oxidation. *Biochim Biophys Acta BBA - Bioenerg* **1827**:94–113.

Cai C, Leu AO, Xie G-J, Guo J, Feng Y, Zhao J-X, Tyson GW, Yuan Z, Hu S. 2018. A methanotrophic archaeon couples anaerobic oxidation of methane to Fe(III) reduction. *ISME J* **12**:1929–1939.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**:421.

Carlson HK, Iavarone AT, Gorur A, Yeo BS, Tran R, Melnyk RA, Mathies RA, Auer M, Coates JD. 2012. Surface multiheme c-type cytochromes from Thermincola potens and implications for respiratory metal reduction by Gram-positive bacteria. *Proc Natl Acad Sci* **109**:1702–1707. doi:10.1073/pnas.1112905109

Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**:1925–1927. doi:10.1093/bioinformatics/btz848

Chen S-C, Musat N, Lechtenfeld OJ, Paschke H, Schmidt M, Said N, Popp D, Calabrese F, Stryhanyuk H, Jaekel U, Zhu Y-G, Joye SB, Richnow H-H, Widdel F, Musat F. 2019. Anaerobic oxidation of ethane by archaea from a marine hydrocarbon seep. *Nature* **568**:108–111. doi:10.1038/s41586-019-1063-0

Chistoserdova L. 1998. C1 Transfer Enzymes and Coenzymes Linking Methylotrophic Bacteria and Methanogenic Archaea. *Science* **281**:99–102. doi:10.1126/science.281.5373.99

Chistoserdova L, Kalyuzhnaya MG, Lidstrom ME. 2009. The Expanding World of Methylotrophic Metabolism. *Annu Rev Microbiol* **63**:477–499. doi:10.1146/annurev.micro.091208.073600

Conley BE, Intile PJ, Bond DR, Gralnick JA. 2018. Divergent Nrf Family Proteins and MtrCAB Homologs Facilitate Extracellular Electron Transfer in *Aeromonas hydrophila*. *Appl Environ Microbiol* **84**. doi:10.1128/AEM.02134-18

Cordova CD, Schicklberger MFR, Yu Y, Spormann AM. 2011. Partial Functional Replacement of CymA by SirCD in Shewanella oneidensis MR-1. *J Bacteriol* **193**:2312–2321. doi:10.1128/JB.01355-10

Costa KC, Lie TJ, Xia Q, Leigh JA. 2013. VhuD Facilitates Electron Flow from H2 or Formate to Heterodisulfide Reductase in Methanococcus maripaludis. *J Bacteriol* **195**:5160–5165. doi:10.1128/JB.00895-13

Costa KC, Wong PM, Wang T, Lie TJ, Dodsworth JA, Swanson I, Burn JA, Hackett M, Leigh JA. 2010. Protein complexing in a methanogen suggests electron bifurcation and electron delivery from formate to heterodisulfide reductase. *Proc Natl Acad Sci U A* **107**:11050–11055.

Dekas AE, Connon SA, Chadwick GL, Trembath-Reichert E, Orphan VJ. 2016. Activity and interactions of methane seep microorganisms assessed by parallel transcription and FISH-NanoSIMS analyses. *ISME J* **10**:678–692. doi:10.1038/ismej.2015.145

Dekas AE, Poretsky RS, Orphan VJ. 2009. Deep-sea archaea fix and share nitrogen in methane-consuming microbial consortia. *Science* **326**:422–426.

Demmer JK, Huang H, Wang S, Demmer U, Thauer RK, Ermler U. 2015. Insights into Flavin-based Electron Bifurcation via the NADH-dependent Reduced Ferredoxin:NADP Oxidoreductase Structure. *J Biol Chem* **290**:21985–21995. doi:10.1074/jbc.M115.656520

Dombrowski N, Seitz KW, Teske AP, Baker BJ. 2017. Genomic insights into potential interdependencies in microbial hydrocarbon and nutrient cycling in hydrothermal sediments. *Microbiome* **5**:106.

Dos Santos PC, Fang Z, Mason SW, Setubal JC, Dixon R. 2012. Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics* **13**:162.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–1797. doi:10.1093/nar/gkh340

Embree M, Liu JK, Al-Bassam MM, Zengler K. 2015. Networks of energetic and metabolic interactions define dynamics in microbial communities. *Proc Natl Acad Sci U A* **112**:15450–15455.

Ettwig KF, Zhu B, Speth D, Keltjens JT, Jetten MSM, Kartal B. 2016. Archaea catalyze iron-dependent anaerobic oxidation of methane. *Proc Natl Acad Sci U A* **113**:12792–12796.

Filman DJ, Marino SF, Ward JE, Yang L, Mester Z, Bullitt E, Lovley DR, Strauss M. 2019. Cryo-EM reveals the structural basis of long-range electron transport in a cytochrome-based bacterial nanowire. *Commun Biol* **2**:219. doi:10.1038/s42003-019-0448-9

Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang H-Y, Dosztányi Z, El-Gebali S, Fraser M, Gough J, Haft D, Holliday GL, Huang H, Huang X, Letunic I, Lopez R, Lu S, Marchler-Bauer A, Mi H, Mistry J, Natale DA, Necci M, Nuka G, Orengo CA, Park Y, Pesseat S, Piovesan D, Potter SC, Rawlings ND, Redaschi N, Richardson L, Rivoire C, Sangrador-Vegas A, Sigrist C, Sillitoe I, Smithers B, Squizzato S, Sutton G, Thanki N, Thomas PD, Tosatto SCE, Wu CH, Xenarios I, Yeh L-S, Young S-Y, Mitchell AL. 2017. InterPro in 2017-beyond protein family and domain annotations. *Nucleic Acids Res* **45**:D190–D199.

Fondi M, Emiliani G, Liò P, Gribaldo S, Fani R. 2009. The Evolution of Histidine Biosynthesis in Archaea: Insights into the his Genes Structure and Organization in LUCA. *J Mol Evol* **69**:512. doi:10.1007/s00239-009-9286-6

Fu H, Metcalf WW. 2015. Genetic Basis for Metabolism of Methylated Sulfur Compounds in Methanosarcina Species. *J Bacteriol* **197**:1515–1524. doi:10.1128/JB.02605-14

Glass JB, Yu H, Steele JA, Dawson KS, Sun S, Chourey K, Pan C, Hettich RL, Orphan VJ. 2014. Geochemical, metagenomic and metaproteomic insights into trace metal utilization by methane-oxidizing microbial consortia in sulphidic marine sediments. *Environ Microbiol* **16**:1592–1611. doi:10.1111/1462-2920.12314

Gong W, Hao B, Wei Z, Ferguson DJ, Tallant T, Krzycki JA, Chan MK. 2008. Structure of the α2ε2 Ni-dependent CO dehydrogenase component of the Methanosarcina barkeri acetyl-CoA decarbonylase/synthase complex. *Proc Natl Acad Sci* **105**:9558–9563. doi:10.1073/pnas.0800415105

Grimaldi S, Schoepp-Cothenet B, Ceccaldi P, Guigliarelli B, Magalon A. 2013. The prokaryotic Mo/W-bisPGD enzymes family: A catalytic workhorse in bioenergetic. *Biochim Biophys Acta BBA - Bioenerg*, Metals in Bioenergetics and Biomimetics Systems **1827**:1048–1085. doi:10.1016/j.bbabio.2013.01.011

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* **59**:307–321. doi:10.1093/sysbio/syq010

Hahn CJ, Laso-Pérez R, Vulcano F, Vaziourakis K-M, Stokke R, Steen IH, Teske A, Boetius A, Liebeke M, Amann R, Knittel K, Wegener G. 2020. "Candidatus Ethanoperedens," a Thermophilic Genus of Archaea Mediating the Anaerobic Oxidation of Ethane. *mBio* **11**. doi:10.1128/mBio.00600-20

Haimovitz R, Barak Y, Morag E, Voronov-Goldman M, Shoham Y, Lamed R, Bayer EA. 2008. Cohesin-dockerin microarray: Diverse specificities between two complementary families of interacting protein modules. *PROTEOMICS* **8**:968–979. doi:10.1002/pmic.200700486

Hallam SJ, Putnam N, Preston CM, Detter JC, Rokhsar D, Richardson PM, DeLong EF. 2004. Reverse Methanogenesis: Testing the Hypothesis with Environmental Genomics. *Science* **305**:1457–1462. doi:10.1126/science.1100025

Hanson RS, Hanson TE. 1996. Methanotrophic bacteria. *Microbiol Rev* **60**:439–471.

Haroon MF, Hu S, Shi Y, Imelfort M, Keller J, Hugenholtz P, Yuan Z, Tyson GW. 2013. Anaerobic oxidation of methane coupled to nitrate reduction in a novel archaeal lineage. *Nature* **500**:567–570.

Hatzenpichler R, Connon SA, Goudeau D, Malmstrom RR, Woyke T, Orphan VJ. 2016. Visualizing in situ translational activity for identifying and sorting slow-growing archaeal-bacterial consortia. *Proc Natl Acad Sci U A* **113**:E4069–78.

Helgadóttir S, Rosas-Sandoval G, Söll D, Graham DE. 2007. Biosynthesis of phosphoserine in the Methanococcales. *J Bacteriol* **189**:575–582.

Hille R, Hall J, Basu P. 2014. The mononuclear molybdenum enzymes. *Chem Rev* **114**:3963–4038.

Hinrichs KU, Hayes JM, Sylva SP, Brewer PG, DeLong EF. 1999. Methane-consuming archaebacteria in marine sediments. *Nature* **398**:802–805.

Hoehler TM, Alperin MJ, Albert DB, Martens CS. 1994. Field and laboratory studies of methane oxidation in an anoxic marine sediment: Evidence for a methanogen-sulfate reducer consortium. *Glob Biogeochem Cycles* **8**:451–463. doi:10.1029/94GB01800

Holler T, Wegener G, Niemann H, Deusner C, Ferdelman TG, Boetius A, Brunner B, Widdel F. 2011. Carbon and sulfur back flux during anaerobic microbial oxidation of methane and coupled sulfate reduction. *Proc Natl Acad Sci* **108**:E1484–E1490. doi:10.1073/pnas.1106032108

Holmes DE, Ueki T, Tang H-Y, Zhou J, Smith JA, Chaput G, Lovley DR. 2019. A Membrane-Bound Cytochrome Enables *Methanosarcina acetivorans* To Conserve Energy from Extracellular Electron Transfer. *mBio* **10**:e00789-19, /mbio/10/4/mBio.00789-19.atom. doi:10.1128/mBio.00789-19

House CH, Orphan VJ, Turk KA, Thomas B, Pernthaler A, Vrentas JM, Joye SB. 2009. Extensive carbon isotopic heterogeneity among methane seep microbiota. *Environ Microbiol* **11**:2207–2215. doi:10.1111/j.1462-2920.2009.01934.x

Hoyt JC, Oren A, Escalante-Semerena JC, Wolfe RS. 1986. Tetrahydromethanopterin-dependent serine transhydroxymethylase from Methanobacterium thermoautotrophicum. *Arch Microbiol* **145**:153–158. doi:10.1007/BF00446773

Hu Y, Ribbe MW. 2016. Biosynthesis of the Metalloclusters of Nitrogenases. *Annu Rev Biochem* **85**:455–483.

Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol Biol Evol* **33**:1635–1638.

Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**:119.

Inoue K, Qian X, Morgado L, Kim B-C, Mester T, Izallalen M, Salgueiro CA, Lovley DR. 2010. Purification and Characterization of OmcZ, an Outer-Surface, Octaheme c-Type Cytochrome Essential for Optimal Current Production by Geobacter sulfurreducens. *Appl Environ Microbiol* **76**:3999–4007. doi:10.1128/AEM.00027-10

Jiménez Otero F, Chan CH, Bond DR. 2018. Identification of Different Putative Outer Membrane Electron Conduits Necessary for Fe(III) Citrate, Fe(III) Oxide, Mn(IV) Oxide, or Electrode Reduction by *Geobacter sulfurreducens*. *J Bacteriol* **200**:e00347-18, /jb/200/19/e00347-18.atom. doi:10.1128/JB.00347-18

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**:1236–1240.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649.

Kellermann MY, Wegener G, Elvert M, Yoshinaga MY, Lin Y-S, Holler T, Mollar XP, Knittel K, Hinrichs K-U. 2012. Autotrophy as a predominant mode of carbon fixation in anaerobic methane-oxidizing microbial communities. *Proc Natl Acad Sci* **109**:19321–19326. doi:10.1073/pnas.1208795109

Kelley DS, Karson JA, Früh-Green GL, Yoerger DR, Shank TM, Butterfield DA, Hayes JM, Schrenk MO, Olson EJ, Proskurowski G, Jakuba M, Bradley A, Larson B, Ludwig K, Glickson D, Buckman K, Bradley AS, Brazelton WJ, Roe K, Elend MJ, Delacour A, Bernasconi SM, Lilley MD, Baross JA, Summons RE, Sylva SP. 2005. A Serpentinite-Hosted Ecosystem: The Lost City Hydrothermal Field. *Science* **307**:1428–1434. doi:10.1126/science.1102556

Kendall MM, Boone DR. 2006. The Order Methanosarcinales In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes: Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes. New York, NY: Springer New York. pp. 244–256. doi:10.1007/0-387-30743-5_12

Kessel M, Volker S, Santarius U, Huber R, Baumeister W. 1990. Three-dimensional Reconstruction of the Surface Protein of the Extremely Thermophilic Archaebacterium Archaeoglobus fulgidus. *Syst Appl Microbiol* **13**:207–213. doi:10.1016/S0723-2020(11)80187-6

Khadem AF, Pol A, Wieczorek A, Mohammadi SS, Francoijs K-J, Stunnenberg HG, Jetten MSM, Op den Camp HJM. 2011. Autotrophic Methanotrophy in Verrucomicrobia: Methylacidiphilum fumariolicumSolV Uses the Calvin-Benson-Bassham Cycle for Carbon Dioxide Fixation. *J Bacteriol* **193**:4438–4446. doi:10.1128/JB.00407-11

Kletzin A, Heimerl T, Flechsler J, van Niftrik L, Rachel R, Klingl A. 2015. Cytochromes c in Archaea: distribution, maturation, cell architecture, and the special case of Ignicoccus hospitalis. *Front Microbiol* **6**. doi:10.3389/fmicb.2015.00439

Knittel K, Boetius A. 2009. Anaerobic oxidation of methane: progress with an unknown process. *Annu Rev Microbiol* **63**:311–334. doi:10.1146/annurev.micro.61.080706.093130

Kojima H, Moll J, Kahnt J, Fukui M, Shima S. 2014. A reversed genetic approach reveals the coenzyme specificity and other catalytic properties of three enzymes putatively involved in anaerobic oxidation of methane with sulfate: Enzymes involved in anaerobic oxidation of methane. *Environ Microbiol* **16**:3431–3442. doi:10.1111/1462-2920.12475

Krüger M, Meyerdierks A, Glöckner FO, Amann R, Widdel F, Kube M, Reinhardt R, Kahnt J, Böcher R, Thauer RK, Shima S. 2003. A conspicuous nickel protein in microbial mats that oxidize methane anaerobically. *Nature* **426**:878–881. doi:10.1038/nature02207

Krukenberg V, Harding K, Richter M, Glöckner FO, Gruber-Vodicka H, Adam B, Berg JS, Knittel K, Tegetmeyer HE, Boetius A, Wegener G. 2016. Candidatus Desulfofervidus auxilii, a hydrogenotrophic sulfate-reducing bacterium involved in the thermophilic anaerobic oxidation of methane. *Env Microbiol*.

Krukenberg V, Riedel D, Gruber-Vodicka HR, Buttigieg PL, Tegetmeyer HE, Boetius A, Wegener G. 2018. Gene expression and ultrastructure of meso- and thermophilic methanotrophic consortia. *Env Microbiol*.

Kulkarni G, Mand TD, Metcalf WW. 2018. Energy Conservation via Hydrogen Cycling in the Methanogenic Archaeon *Methanosarcina barkeri*. *mBio* **9**. doi:10.1128/mBio.01256-18

Lane DJ. 1991. 16S/23S rRNA sequencing In: Stackebrandt E, Goodfellow M, editors. Nucleic Acid Techniques in Bacterial Systematics. New York, NY: John Wiley and Sons. pp. 115–147.

Laso-Pérez R, Wegener G, Knittel K, Widdel F, Harding KJ, Krukenberg V, Meier DV, Richter M, Tegetmeyer HE, Riedel D, Richnow H-H, Adrian L, Reemtsma T, Lechtenfeld O, Musat F. 2016. Thermophilic archaea activate butane via alkyl-coenzyme M formation. *Nature*.

Latka A, Maciejewska B, Majkowska-Skrobek G, Briers Y, Drulis-Kawa Z. 2017. Bacteriophage-encoded virion-associated enzymes to overcome the carbohydrate barriers during the infection process. *Appl Microbiol Biotechnol* **101**:3103–3119.

Leadbetter ER, Foster JW. 1958. Studies on Some Methane-Utilizing Bacteria. *Arch Microbiol* **30**:91–118.

Lechner M, Findeiß S, Steiner L, Marz M, Stadler P, Prohaska S. 2011. Proteinortho: Detection of (Co-)orthologs in large-scale analysis. *BMC Bioinformatics* **12**:124.

Lee J-H, Harvat EM, Stevens JM, Ferguson SJ, Saier MH. 2007. Evolutionary origins of members of a superfamily of integral membrane cytochrome c biogenesis proteins. *Biochim Biophys Acta BBA - Biomembr* **1768**:2164–2181. doi:10.1016/j.bbamem.2007.04.022

Lefort V, Longueville J-E, Gascuel O. 2017. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* **34**:2422–2424. doi:10.1093/molbev/msx149

Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res* **47**:W256–W259. doi:10.1093/nar/gkz239

Leu AO, McIlroy SJ, Ye J, Parks DH, Orphan VJ, Tyson GW. 2020. Lateral gene transfer drives metabolic flexibility in the anaerobic methane oxidising archaeal family Methanoperedenaceae. *bioRxiv* 2020.02.06.936641. doi:10.1101/2020.02.06.936641

Levar CE, Chan CH, Mehta-Kolte MG, Bond DR. 2014. An Inner Membrane Cytochrome Required Only for Reduction of High Redox Potential Extracellular Electron Acceptors. *mBio* **5**:e02034-14. doi:10.1128/mBio.02034-14

Li F, Hinderberger J, Seedorf H, Zhang J, Buckel W, Thauer RK. 2008. Coupled Ferredoxin and Crotonyl Coenzyme A (CoA) Reduction with NADH Catalyzed by the Butyryl-CoA Dehydrogenase/Etf Complex from Clostridium kluyveri. *J Bacteriol* **190**:843–850. doi:10.1128/JB.01417-07

Li Q, Li L, Rejtar T, Lessner DJ, Karger BL, Ferry JG. 2006. Electron transport in the pathway of acetate conversion to methane in the marine archaeon Methanosarcina acetivorans. *J Bacteriol* **188**:702–710.

Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**:1658–1659.

Lidstrom ME. 2006. Aerobic Methylotrophic Prokaryotes In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes: Volume 2: Ecophysiology and Biochemistry. New York, NY: Springer. pp. 618–634. doi:10.1007/0-387-30742-7_20

Lin Z, Sparling R. 1998. Investigation of serine hydroxymethyltransferase in methanogens. *Can J Microbiol* **44**:652–656. doi:10.1139/w98-050

Lloyd KG, Alperin MJ, Teske A. 2011. Environmental evidence for net methane production and oxidation in putative ANaerobic MEthanotrophic (ANME) archaea. *Environ Microbiol* **13**:2548–2564. doi:10.1111/j.1462-2920.2011.02526.x

Lösekann T, Knittel K, Nadalig T, Fuchs B, Niemann H, Boetius A, Amann R. 2007. Diversity and abundance of aerobic and anaerobic methane oxidizers at the Haakon Mosby Mud Volcano, Barents Sea. *Appl Env Microbiol* **73**:3348–3362.

Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, Buchner A, Lai T, Steppi S, Jobb G, F$\surd$$\partial$rster W, Brettske I, Gerber S, Ginhart AW, Gross O, Grumann S, Hermann S, Jost R, K$\surd$$\partial$nig A, Liss T, L$\surd$°$\surd$ümann R, May M, Nonhoff B, Reichel B, Strehlow R, Stamatakis A, Stuckmann N, Vilbig A, Lenke M, Ludwig T, Bode A, Schleifer K. 2004. ARB: a software environment for sequence data. *Nucleic Acids Res* **32**:1363–1371.

Mamat B, Roth A, Grimm C, Ermler U, Tziatzios C, Schubert D, Thauer RK, Shima S. 2009. Crystal structures and enzymatic properties of three formyltransferases from archaea: Environmental adaptation and evolutionary relationship. *Protein Sci* **11**:2168–2178. doi:10.1110/ps.0211002

McGlynn SE. 2017. Energy Metabolism during Anaerobic Methane Oxidation in ANME Archaea. *Microbes Env* **32**:5–13.

McGlynn SE, Boyd ES, Peters JW, Orphan VJ. 2012. Classifying the metal dependence of uncharacterized nitrogenases. *Front Microbiol* **3**:419.

McGlynn SE, Chadwick GL, Kempes CP, Orphan VJ. 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* **526**:531–535.

McGlynn SE, Chadwick GL, O'Neill A, Mackey M, Thor A, Deerinck TJ, Ellisman MH, Orphan VJ. 2018. Subgroup characteristics of marine methane-oxidizing ANME-2 archaea and their syntrophic partners revealed by integrated multimodal analytical microscopy. *Appl Env Microbiol*.

Meulepas RJW, Jagersma CG, Khadem AF, Stams AJM, Lens PNL. 2010. Effect of methanogenic substrates on anaerobic oxidation of methane and sulfate reduction by an anaerobic methanotrophic enrichment. *Appl Microbiol Biotechnol* **87**:1499–1506. doi:10.1007/s00253-010-2597-0

Mevers E, Su L, Pishchany G, Baruch M, Cornejo J, Hobert E, Dimise E, Ajo-Franklin CM, Clardy J. 2019. An elusive electron shuttle from a facultative anaerobe. *eLife* **8**:e48054. doi:10.7554/eLife.48054

Meyerdierks A, Kube M, Kostadinov I, Teeling H, Glöckner FO, Reinhardt R, Amann R. 2010. Metagenome and mRNA expression analyses of anaerobic methanotrophic archaea of the ANME-1 group. *Env Microbiol* **12**:422–439.

Meyerdierks A, Kube M, Lombardot T, Knittel K, Bauer M, Glöckner FO, Reinhardt R, Amann R. 2005. Insights into the genomes of archaea mediating the anaerobic oxidation of methane. *Environ Microbiol* **7**:1937–1951. doi:10.1111/j.1462-2920.2005.00844.x

Mills DJ, Vitt S, Strauss M, Shima S, Vonck J. 2013. De novo modeling of the F(420)-reducing [NiFe]-hydrogenase from a methanogenic archaeon by cryo-electron microscopy. *Elife* **2**:e00218.

Milton RD, Ruth JC, Deutzmann JS, Spormann AM. 2018. Methanococcus maripaludis Employs Three Functional Heterodisulfide Reductase Complexes for Flavin-Based Electron Bifurcation Using Hydrogen and Formate. *Biochemistry*. doi:10.1021/acs.biochem.8b00662

Milucka J, Ferdelman TG, Polerecky L, Franzke D, Wegener G, Schmid M, Lieberwirth I, Wagner M, Widdel F, Kuypers MMM. 2012. Zero-valent sulphur is a key intermediate in marine methane oxidation. *Nature* **491**:541–546.

Miyazaki J, Higa R, Toki T, Ashi J, Tsunogai U, Nunoura T, Imachi H, Takai K. 2009. Molecular Characterization of Potential Nitrogen Fixation by Anaerobic Methane-Oxidizing Archaea in the Methane Seep Sediments at the Number 8 Kumano Knoll in the Kumano Basin, Offshore of Japan. *Appl Env Microbiol* **75**:7153–7162. doi:10.1128/AEM.01184-09

Mock J, Wang S, Huang H, Kahnt J, Thauer RK. 2014. Evidence for a Hexaheteromeric Methylenetetrahydrofolate Reductase in Moorella thermoacetica. *J Bacteriol* **196**:3303–3314. doi:10.1128/JB.01839-14

Moran JJ, Beal EJ, Vrentas JM, Orphan VJ, Freeman KH, House CH. 2008. Methyl sulfides as intermediates in the anaerobic oxidation of methane. *Env Microbiol* **10**:162–173.
Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. *mBio* **3**. doi:10.1128/mBio.00036-12
Myllykallio H. 2002. An Alternative Flavin-Dependent Mechanism for Thymidylate Synthesis. *Science* **297**:105–107. doi:10.1126/science.1072113
Nauhaus K, Albrecht M, Elvert M, Boetius A, Widdel F. 2007. In vitro cell growth of marine archaeal-bacterial consortia during anaerobic oxidation of methane with sulfate. *Env Microbiol* **9**:187–196.
Nauhaus K, Boetius A, Krüger M, Widdel F. 2002. In vitro demonstration of anaerobic oxidation of methane coupled to sulphate reduction in sediment from a marine gas hydrate area. *Env Microbiol* **4**:296–305.
Nauhaus K, Treude T, Boetius A, Krüger M. 2005. Environmental regulation of the anaerobic oxidation of methane: a comparison of ANME-I and ANME-II communities. *Env Microbiol* **7**:98–106.
Newman DK, Kolter R. 2000. A role for excreted quinones in extracellular electron transfer. *Nature* **405**:94–97. doi:10.1038/35011098
Niemann H, Losekann T, de Beer D, Elvert M, Nadalig T, Knittel K, Amann R, Sauter EJ, Schluter M, Klages M, Foucher JP, Boetius A. 2006. Novel microbial communities of the Haakon Mosby mud volcano and their role as a methane sink. *Nature* **443**:854–858. doi:10.1038/nature05227
Orphan VJ, Hinrichs K-U, Ussler W, Paull CK, Taylor LT, Sylva SP, Hayes JM, Delong EF. 2001a. Comparative Analysis of Methane-Oxidizing Archaea and Sulfate-Reducing Bacteria in Anoxic Marine Sediments. *Appl Env Microbiol* **67**:1922–1934. doi:10.1128/AEM.67.4.1922-1934.2001
Orphan VJ, House CH, Hinrichs K-U, McKeegan KD, DeLong EF. 2002. Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proc Natl Acad Sci* **99**:7663–7668. doi:10.1073/pnas.072210299
Orphan VJ, House CH, Hinrichs K-U, McKeegan KD, DeLong EF. 2001b. Methane-Consuming Archaea Revealed by Directly Coupled Isotopic and Phylogenetic Analysis. *Science* **293**:484–487.
Pancost RD, Damsté JSS, Lint S de, Maarel MJEC van der, Gottschal JC, Party‡ TMSS. 2000. Biomarker Evidence for Widespread Anaerobic Methane Oxidation in Mediterranean Sediments by a Consortium of Methanogenic Archaea and Bacteria. *Appl Environ Microbiol* **66**:1126–1132. doi:10.1128/AEM.66.3.1126-1132.2000
Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**:996–1004. doi:10.1038/nbt.4229
Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**:1043–1055.
Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol*.
Paul L, Ferguson DJ, Krzycki JA. 2000. The Trimethylamine Methyltransferase Gene and Multiple Dimethylamine Methyltransferase Genes of Methanosarcina barkeri Contain In-Frame and Read-Through Amber Codons. *J Bacteriol* **182**:2520–2529. doi:10.1128/JB.182.9.2520-2529.2000
Peer A, Smith SP, Bayer EA, Lamed R, Borovok I. 2009. Noncellulosomal cohesin- and dockerin-like modules in the three domains of life. *FEMS Microbiol Lett* **291**:1–16. doi:10.1111/j.1574-6968.2008.01420.x
Pernthaler A, Dekas AE, Brown CT, Goffredi SK, Embaye T, Orphan VJ. 2008. Diverse syntrophic partnerships from deep-sea methane vents revealed by direct cell capture and metagenomics. *Proc Natl Acad Sci* **105**:7052–7057. doi:10.1073/pnas.0711303105
Porat I, Sieprawska-Lupa M, Teng Q, Bohanon FJ, White RH, Whitman WB. 2006. Biochemical and genetic characterization of an early step in a novel pathway for the biosynthesis of aromatic amino acids and p-aminobenzoic acid in the archaeon Methanococcus maripaludis. *Mol Microbiol* **62**:1117–1131.

Prakash D, Chauhan SS, Ferry JG. 2019. Life on the thermodynamic edge: Respiratory growth of an acetotrophic methanogen. *Sci Adv* **5**:eaaw9059. doi:10.1126/sciadv.aaw9059

Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**:1823–1829.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Gl$\surd$$\partial$ckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**:D590–D596.

Raemakers-Franken PC, Bongaerts R, Fokkens R, Drift C, Vogels GD. 1991. Characterization of two pterin derivatives isolated from Methanoculleus thermophilicum. *Eur J Biochem* **200**:783–787. doi:10.1111/j.1432-1033.1991.tb16245.x

Rasigraf O, Kool DM, Jetten MSM, Sinninghe Damsté JS, Ettwig KF. 2014. Autotrophic Carbon Dioxide Fixation via the Calvin-Benson-Bassham Cycle by the Denitrifying Methanotroph "Candidatus Methylomirabilis oxyfera." *Appl Environ Microbiol* **80**:2451–2460. doi:10.1128/AEM.04199-13

Richardson DJ, Butt JN, Fredrickson JK, Zachara JM, Shi L, Edwards MJ, White G, Baiden N, Gates AJ, Marritt SJ, Clarke TA. 2012. The "porin-cytochrome" model for microbe-to-mineral electron transfer. *Mol Microbiol* **85**:201–212.

Rohlin L, Gunsalus RP. 2010. Carbon-dependent control of electron transfer and central carbon pathway genes for methane biosynthesis in the Archaean, Methanosarcina acetivorans strain C2A. *BMC Microbiol* **10**:62. doi:10.1186/1471-2180-10-62

Sanders C, Turkarslan S, Lee D-W, Daldal F. 2010. Cytochrome c biogenesis: the Ccm system. *Trends Microbiol* **18**:266–274. doi:10.1016/j.tim.2010.03.006

Sarris PF, Ladoukakis ED, Panopoulos NJ, Scoulica EV. 2014. A Phage Tail-Derived Element with Wide Distribution among Both Prokaryotic Domains: A Comparative Genomic and Phylogenetic Study. *Genome Biol Evol* **6**:1739–1747. doi:10.1093/gbe/evu136

Sauer K, Harms U, Thauer RK. 1997. Methanol: Coenzyme M Methyltransferase from Methanosarcina Barkeri. Purification, Properties and Encoding Genes of the Corrinoid Protein MT1. *Eur J Biochem* **243**:670–677. doi:10.1111/j.1432-1033.1997.t01-1-00670.x

Scheller S, Goenrich M, Boecher R, Thauer RK, Jaun B. 2010. The key nickel enzyme of methanogenesis catalyses the anaerobic oxidation of methane. *Nature* **465**:606–608.

Scheller S, Yu H, Chadwick GL, McGlynn SE, Orphan VJ. 2016. Artificial electron acceptors decouple archaeal methane oxidation from sulfate reduction. *Science* **351**:703–707.

Schink B. 1997. Energetics of syntrophic cooperation in methanogenic degradation. *Microbiol Mol Biol Rev* **61**:262–280.

Schink B, Stams AJM. 2006. Syntrophism among Prokaryotes In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes. New York, NY: Springer New York. pp. 309–335. doi:10.1007/0-387-30742-7_11

Schlegel K, Welte C, Deppenmeier U, Müller V. 2012. Electron transport during aceticlastic methanogenesis by Methanosarcina acetivorans involves a sodium-translocating Rnf complex. *FEBS J* **279**:4444–4452.

Schroder I, Thauer RK. 1999. Methylcobalamin:homocysteine methyltransferase from Methanobacterium thermoautotrophicum . Identification as the metE gene product. *Eur J Biochem* **263**:789–796. doi:10.1046/j.1432-1327.1999.00559.x

Schwalb C, Chapman SK, Reid GA. 2003. The Tetraheme Cytochrome CymA Is Required for Anaerobic Respiration with Dimethyl Sulfoxide and Nitrite in *Shewanella oneidensis* †. *Biochemistry* **42**:9491–9497. doi:10.1021/bi034456f

Shi L, Dong H, Reguera G, Beyenal H, Lu A, Liu J, Yu H-Q, Fredrickson JK. 2016. Extracellular electron transfer mechanisms between microorganisms and minerals. *Nat Rev Microbiol* **14**:651–662. doi:10.1038/nrmicro.2016.93

Shikuma NJ, Pilhofer M, Weiss GL, Hadfield MG, Jensen GJ, Newman DK. 2014. Marine Tubeworm Metamorphosis Induced by Arrays of Bacterial Phage Tail-Like Structures. *Science* **343**:529–533. doi:10.1126/science.1246794

Shima S, Krueger M, Weinert T, Demmer U, Kahnt J, Thauer RK, Ermler U. 2012. Structure of a methyl-coenzyme M reductase from Black Sea mats that oxidize methane anaerobically. *Nature* **481**:98–101. doi:10.1038/nature10663

Shima S, Warkentin E, Grabarse W, Sordel M, Wicke M, Thauer RK, Ermler U. 2000. Structure of coenzyme F(420) dependent methylenetetrahydromethanopterin reductase from two methanogenic archaea. *J Mol Biol* **300**:935–950.

Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG. 2013. PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. *Nature* **500**:350–353. doi:10.1038/nature12453

Simon J, van Spanning RJM, Richardson DJ. 2008. The organisation of proton motive and non-proton motive redox loops in prokaryotic respiratory systems. *Biochim Biophys Acta BBA - Bioenerg* **1777**:1480–1490. doi:10.1016/j.bbabio.2008.09.008

Simpson PG, Whitman WB. 1993. Anabolic Pathways in Methanogens In: Ferry JG, editor. Methanogenesis: Ecology, Physiology, Biochemistry & Genetics. Boston, MA: Springer US. pp. 445–472. doi:10.1007/978-1-4615-2391-8_11

Skennerton CT, Chourey K, Iyer R, Hettich RL, Tyson GW, Orphan VJ. 2017. Methane-Fueled Syntrophy through Extracellular Electron Transfer: Uncovering the Genomic Traits Conserved within Diverse Bacterial Partners of Anaerobic Methanotrophic Archaea. *MBio* **8**.

Slobodkina GB, Kolganova TV, Querellou J, Bonch-Osmolovskaya EA, Slobodkin AI. 2009. Geoglobus acetivorans sp. nov., an iron(III)-reducing archaeon from a deep-sea hydrothermal vent. *Int J Syst Evol Microbiol* **59**:2880–2883.

Smith JA, Aklujkar M, Risso C, Leang C, Giloteaux L, Holmes DE. 2015. Mechanisms involved in Fe(III) respiration by the hyperthermophilic archaeon Ferroglobus placidus. *Appl Env Microbiol* **81**:2735–2744.

Sørensen KB, Finster K, Ramsing NB. 2001. Thermodynamic and kinetic requirements in anaerobic methane oxidizing consortia exclude hydrogen, acetate, and methanol as possible electron shuttles. *Microb Ecol* **42**:1–10. doi:10.1007/s002480000083

Spang A, Caceres EF, Ettema TJG. 2017. Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* **357**:eaaf3883. doi:10.1126/science.aaf3883

Sparacino-Watkins C, Stolz JF, Basu P. 2014. Nitrate and periplasmic nitrate reductases. *Chem Soc Rev* **43**:676–706. doi:10.1039/C3CS60249D

Speare L, Cecere AG, Guckes KR, Smith S, Wollenberg MS, Mandel MJ, Miyashiro T, Septer AN. 2018. Bacterial symbionts use a type VI secretion system to eliminate competitors in their natural host. *Proc Natl Acad Sci* **115**:E8528–E8537. doi:10.1073/pnas.1808302115

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313.

Staples CR, Lahiri S, Raymond J, Von Herbulis L, Mukhophadhyay B, Blankenship RE. 2007. Expression and Association of Group IV Nitrogenase NifD and NifH Homologs in the Non-Nitrogen-Fixing Archaeon Methanocaldococcus jannaschii. *J Bacteriol* **189**:7392–7398. doi:10.1128/JB.00876-07

Stojanowic A, Mander GJ, Duin EC, Hedderich R. 2003. Physiological role of the F420-non-reducing hydrogenase (Mvh) from Methanothermobacter marburgensis. *Arch Microbiol* **180**:194–203.

Stokke R, Roalkvam I, Lanzen A, Haflidason H, Steen IH. 2012. Integrated metagenomic and metaproteomic analyses of an ANME-1-dominated community in marine cold seep sediments. *Env Microbiol* **14**:1333–1346.

Susanti D, Mukhopadhyay B. 2012. An Intertwined Evolutionary History of Methanogenic Archaea and Sulfate Reduction. *PLOS ONE* **7**:e45313. doi:10.1371/journal.pone.0045313

Takeuchi M, Yoshioka H, Seo Y, Tanabe S, Tamaki H, Kamagata Y, Takahashi HA, Igari S, Mayumi D, Sakata S. 2011. A distinct freshwater-adapted subgroup of ANME-1 dominates active archaeal

communities in terrestrial subsurfaces in Japan. *Environ Microbiol* **13**:3206–3218. doi:10.1111/j.1462-2920.2011.02517.x

Taylor NMI, Raaij MJ van, Leiman PG. 2018. Contractile injection systems of bacteriophages and related systems. *Mol Microbiol* **108**:6–15. doi:10.1111/mmi.13921

Thauer RK. 2011. Anaerobic oxidation of methane with sulfate: on the reversibility of the reactions that are catalyzed by enzymes also involved in methanogenesis from CO2. *Curr Opin Microbiol* **14**:292–299. doi:10.1016/j.mib.2011.03.003

Thauer RK, Kaster A-K, Goenrich M, Schick M, Hiromoto T, Shima S. 2010. Hydrogenases from methanogenic archaea, nickel, a novel cofactor, and H2 storage. *Annu Rev Biochem* **79**:507–536.

Timmers PHA, Welte CU, Koehorst JJ, Plugge CM, Jetten MSM, Stams AJM. 2017. Reverse Methanogenesis and Respiration in Methanotrophic Archaea. *Archaea* **2017**.

Trembath-Reichert E, Case DH, Orphan VJ. 2016. Characterization of microbial associations with methanotrophic archaea and sulfate-reducing bacteria through statistical comparison of nested Magneto-FISH enrichments. *PeerJ* **4**:e1913.

Treude T, Orphan V, Knittel K, Gieseke A, House CH, Boetius A. 2007. Consumption of Methane and CO2 by Methanotrophic Microbial Mats from Gas Seeps of the Anoxic Black Sea. *Appl Environ Microbiol* **73**:2271–2283. doi:10.1128/AEM.02685-06

Troselj V, Treuner-Lange A, Søgaard-Andersen L, Wall D. 2018. Physiological Heterogeneity Triggers Sibling Conflict Mediated by the Type VI Secretion System in an Aggregative Multicellular Bacterium. *mBio* **9**. doi:10.1128/mBio.01645-17

Valentine DL, Reeburgh WS. 2000. New perspectives on anaerobic methane oxidation. *Env Microbiol* **2**:477–484.

von Wintzingerode F, Selent B, Hegemann W, Göbel UB. 1999. Phylogenetic analysis of an anaerobic, trichlorobenzene-transforming microbial consortium. *Appl Env Microbiol* **65**:283–286.

Vorholt JA, Chistoserdova L, Lidstrom ME, Thauer RK. 1998. The NADP-Dependent Methylene Tetrahydromethanopterin Dehydrogenase in Methylobacterium extorquens AM1. *J Bacteriol* **180**:5351–5356.

Wagner T, Koch J, Ermler U, Shima S. 2017. Methanogenic heterodisulfide reductase (HdrABC-MvhAGD) uses two noncubane [4Fe-4S] clusters for reduction. *Science* **357**:699–703. doi:10.1126/science.aan0425

Wang F, Gu Y, O'Brien JP, Yi SM, Yalcin SE, Srikanth V, Shen C, Vu D, Ing NL, Hochbaum AI, Egelman EH, Malvankar NS. 2019. Structure of Microbial Nanowires Reveals Stacked Hemes that Transport Electrons over Micrometers. *Cell* **177**:361-369.e10. doi:10.1016/j.cell.2019.03.029

Wang F-P, Zhang Y, Chen Y, He Y, Qi J, Hinrichs K-U, Zhang X-X, Xiao X, Boon N. 2013. Methanotrophic archaea possessing diverging methane-oxidizing and electron-transporting pathways. *ISME J*.

Wang M, Tomb J-F, Ferry JG. 2011. Electron transport in acetate-grown Methanosarcina acetivorans. *BMC Microbiol* **11**:165. doi:10.1186/1471-2180-11-165

Wang Y, Huang Y, Wang J, Cheng C, Huang W, Lu P, Xu Y-N, Wang P, Yan N, Shi Y. 2009. Structure of the formate transporter FocA reveals a pentameric aquaporin-like channel. *Nature* **462**:467–472. doi:10.1038/nature08610

Wegener G, Krukenberg V, Riedel D, Tegetmeyer HE, Boetius A. 2015. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**:587–590.

Wegener G, Krukenberg V, Ruff SE, Kellermann MY, Knittel K. 2016. Metabolic Capabilities of Microorganisms Involved in and Associated with the Anaerobic Oxidation of Methane. *Front Microbiol* **7**:46.

Wegener G, Niemann H, Elvert M, Hinrichs K-U, Boetius A. 2008. Assimilation of methane and inorganic carbon by microbial communities mediating the anaerobic oxidation of methane. *Environ Microbiol* **10**:2287–2298. doi:10.1111/j.1462-2920.2008.01653.x

Weimer PJ, Zeikus JG. 1978. One carbon metabolism in methanogenic bacteria. *Arch Microbiol* **119**:49–57. doi:10.1007/BF00407927

Welte C, Deppenmeier U. 2014. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochim Biophys Acta BBA - Bioenerg*, 18th European Bioenergetics Conference 2014 Lisbon, Portugal **1837**:1130–1147. doi:10.1016/j.bbabio.2013.12.002

Welte C, Deppenmeier U. 2011a. Membrane-Bound Electron Transport in Methanosaeta thermophila. *J Bacteriol* **193**:2868–2870. doi:10.1128/JB.00162-11

Welte C, Deppenmeier U. 2011b. Re-evaluation of the function of the F420 dehydrogenase in electron transport of Methanosarcina mazei: Role of F420 dehydrogenase in Ms. mazei. *FEBS J* **278**:1277–1287. doi:10.1111/j.1742-4658.2011.08048.x

White RH. 2004. L-Aspartate semialdehyde and a 6-deoxy-5-ketohexose 1-phosphate are the precursors to the aromatic amino acids in Methanocaldococcus jannaschii. *Biochemistry* **43**:7618–7627.

White RH, Xu H. 2006. Methylglyoxal is an intermediate in the biosynthesis of 6-deoxy-5-ketofructose-1-phosphate: a precursor for aromatic amino acid biosynthesis in Methanocaldococcus jannaschii. *Biochemistry* **45**:12366–12379.

Wickham H. 2009. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wood GE, Haydock AK, Leigh JA. 2003. Function and Regulation of the Formate Dehydrogenase Genes of the Methanogenic Archaeon Methanococcus maripaludis. *J Bacteriol* **185**:2548–2554. doi:10.1128/JB.185.8.2548-2554.2003

Yan Z, Ferry JG. 2018. Electron Bifurcation and Confurcation in Methanogenesis and Reverse Methanogenesis. *Front Microbiol* **9**. doi:10.3389/fmicb.2018.01322

Yan Z, Wang M, Ferry JG. 2017. A Ferredoxin- and F420H2-Dependent, Electron-Bifurcating, Heterodisulfide Reductase with Homologs in the Domains Bacteria and Archaea. *MBio* **8**.

Yee MO, Rotaru A-E. 2020. Extracellular electron uptake in Methanosarcinales is independent of multiheme c -type cytochromes. *Sci Rep* **10**:1–12. doi:10.1038/s41598-019-57206-z

Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, Schweer T, Peplies J, Ludwig W, Glöckner FO. 2014. The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* **42**:D643–8.

Yu H, Susanti D, McGlynn SE, Skennerton CT, Chourey K, Iyer R, Scheller S, Tavormina PL, Hettich RL, Mukhopadhyay B, Orphan VJ. 2018. Comparative Genomics and Proteomic Analysis of Assimilatory Sulfate Reduction Pathways in Anaerobic Methanotrophic Archaea. *Front Microbiol* **9**. doi:10.3389/fmicb.2018.02917

Zacharoff L, Chan CH, Bond DR. 2016. Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of Geobacter sulfurreducens. *Bioelectrochemistry* **107**:7–13. doi:10.1016/j.bioelechem.2015.08.003

Zehnder AJB, Brock TD. 1980. Anaerobic Methane Oxidation: Occurrence and Ecology. *Appl Environ Microbiol* **39**:194–204.

Zeikus JG, Fuchs G, Kenealy W, Thauer RK. 1977. Oxidoreductases Involved in Cell Carbon Synthesis of Methanobacterium thermoautotrophicum. *J Bacteriol* **132**:604–613.

Zheng B, Zhang Q, Gao J, Han H, Li M, Zhang J, Qi J, Yan J, Gao GF. 2011. Insight into the Interaction of Metal Ions with TroA from Streptococcus suis. *PLoS ONE* **6**:e19510. doi:10.1371/journal.pone.0019510

Zheng H, Dietrich C, Radek R, Brune A. 2016. Endomicrobium proavitum, the first isolate of Endomicrobia class. nov. (phylum Elusimicrobia)–an ultramicrobacterium with an unusual cell cycle that fixes nitrogen with a Group IV nitrogenase. *Env Microbiol* **18**:191–204.

Chapter 4

4. RAPID BACTERIAL SWIMMING IN "*CANDIDATUS* THIORHODOSPHAERA CELERRIMA", A NOVEL GENUS OF PURPLE SULFUR BACTERIA

Grayson L Chadwick[1*], Christof Meile[2], Kurt W. Hanselmann[3], Dianne K Newmann[1,4], Jared R Leadbetter[1,4]

[1]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA.
[2]Department of Marine Sciences, University of Georgia, Athens, GA, USA.
[3]Department of Earth Sciences, ETH, Zürich, Switzerland
[4]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA.

Abstract

Substrate uptake by bacteria is a diffusion-dominated process. The speed with which most bacteria are able to swim is not great enough to increase substrate uptake by fluid flow, and therefore motility is largely used to find locations with higher substrate concentrations, which leads to greater rates of diffusion. A notable exception to this rule is *Thiovulum majus*, a chemotrophic sulfide oxidizer 10 μm in diameter that can swim up to 600 μm/s. This prodigious swimming speed allows veil-like biofilms of *T. majus* found on sulfidic sediments to significantly increase their oxygen uptake through advective fluid flow. Here we enrich and characterize a representative of a novel genus of purple sulfur bacteria with similar size and swimming speed as *T. majus*. Phylogenetic and genomic analysis places this novel phototroph in the family *Chromatiaceae*, and high-speed video microscopy demonstrates a maximum swimming speed greater than 800 μm/s, making these organisms the second fastest bacterium described. Modeling results suggest that sediment-adhered cells could use this motility to increase the supply of sulfide, the electron donor for anoxygenic photosynthesis, extending this diffusion-beating strategy to a new guild of microorganisms.

Introduction

Sulfide oxidizing bacteria are found in narrow horizons corresponding to optimal combinations of substrates that have steep, opposing concentration gradients. Sulfide is generally produced at depth in anoxic sediments and diffuses upwards into more oxic overlying water where it is consumed by biotic and abiotic processes. Chemotrophic sulfide oxidizers gain energy from the oxidation of sulfide and reduction of higher potential electron acceptors such as nitrate and oxygen, while phototrophic sulfide oxidizers use sulfide as a source of electrons for light-driven cyclic electron flow. Chemotrophic sulfur oxidizing bacteria are referred to as "colorless sulfur bacteria" in contrast to the phototrophic sulfur oxidizers that are split into "purple sulfur bacteria" and "green sulfur bacteria" depending on their phylogeny and complement of photosynthetic pigments. Both modes of growth are generally restricted to upper layers of sulfidic sediments or near the oxycline in euxinic bodies of water, where optimal concentrations of sulfide, electron acceptor and light are achieved. The competition between and within these groups for common resources has been the subject of much investigation, through environmental surveys (Jørgensen and Des Marais, 1986; Visscher et al., 1992), coculture experiments (van Gemerden, 1974; van Gemerden and Beeftink, 1981), and modeling studies (Decker et al., 2005; Wit et al., 1995). Ecophysiological considerations lead to the vertical stratification of sulfide oxidizers in narrow bands corresponding to their optimal photochemical niche.

While nearly all bacteria exist in a low Reynolds number regime where nutrient delivery is dominated by diffusion (Purcell, 1977), certain colorless sulfur bacteria have adaptations thought to circumvent these limitations (Schulz and Jørgensen, 2001). One major mechanism of to achieve substrate uptake rates greater than diffusion is the formation of extraordinarily large cells coupled to physical movement across their opposing substrate gradients. Members of the *Thioploca* genus form giant filaments up to 40 μm

wide and 10s of mm long that migrate vertically between sulfidic sediments and nitrate rich surface waters, filling enormous vacuoles with nitrate at concentrations up to 500 mM (Fossing et al., 1995; Huettel et al., 1996). Similarly sized *Beggiatoa* species form large mats on sulfidic sediments and individual filaments have been observed to react to light and oxygen in order to optimally position themselves in response to sulfide gradients (Møller et al., 1985). *Thiomargarita namibiensis*, the largest bacteria by volume, is mostly composed of a single giant vacuole that is thought to refill with nitrate when they are occasionally resuspended from sulfidic sediment into the overlying water by bubbles created from methane degassing (Schulz and Beer, 2002).

Another mechanism for colorless sulfur bacteria to escape diffusion limitation is to move exceptionally fast. The best studied example of this is found in the genus *Thiovulum*. At approximately 10 μm in diameter, *Thiovulum majus* is smaller than the colorless sulfur bacteria described above, but individual cells can swim with speeds up to 600 μm/s (Garcia-Pichel, 1989). While rapidly motile as free-swimming cells, *Thiovulum* sp. are naturally found in biofilms, known as "veils", comprised of huge numbers of cells anchored to the sediment and one another via "slime" tethers (Fenchel and Glud, 1998). Instead of using their motility in a traditional chemotactic manner to find locations with more optimal substrate concentrations through a biased random walk, *Thiovulum* sp. pull aerobic water towards themselves through the concerted efforts of the biofilm community (Fenchel and Glud, 1998; Petroff et al., 2015; Schulz and Jørgensen, 2001). The size and ultrafast swimming places *Thiovulum* sp. near the boundary between advection-dominated and diffusion-dominated regimes, where substrate uptake can be significantly increased by fluid flow.

In contrast to the colorless sulfur bacteria, the purple sulfur bacteria have not been reported to have similar ability to increase their flux of substrate beyond what diffusion provides. Purple sulfur bacteria are not exceptionally large, precluding the size-dependent methods described above, and while they exhibit various forms of taxis in response to light and chemical stimuli (Gest, 1995; Thar and Kühl, 2001), these are not particularly rapid; they are thought to be used to position themselves at an optimal horizon in sediment or water columns through traditional chemo- and phototactic behavior.

Here we present the physical enrichment and characterization of a novel purple sulfur bacterium that shares some of the unique behaviors of *Thiovulum* sp. Multiple enrichments from two separate marshlands produced morphologically, phylogenetically, and behaviorally consistent populations. Phylogenetic analysis of marker genes, comparative genomics, and whole-cell absorbance spectra supported the placement of this group within the family *Chromatiaceae*. Enriched cells were significantly larger than most other known *Chromatiaceae*, and high-speed video microscopy was used to quantify swimming speeds, revealing the fastest phototrophic bacterium known, with top speeds exceeding those reported for *Thiovulum majus*. We propose the provisional name "*Candidatus* Thiorhodosphaera celerrima" to describe a novel genus in recognition of its phylogenetic distance from existing members of the *Chromatiaceae* as well as the phenotypic peculiarity of their extreme swimming speed (Murray and Stackebrandt, 1995).

Results

Enrichment and morphological description of a novel purple bacteria

Observation of marine-impacted marshlands near Woods Hole, Ma. revealed a thin purple layer overlying sulfidic sediments rich in organic detritus (**Figure 1a, b**). In some places this purple layer was

covered by presumed oxygenic phototrophs based on color and the entrainment of small bubbles. This colored layer and ~7.5 ml of underlying sediment was collected in 50 ml polypropylene tubes, filled with overlying water, capped, and taken to an east facing windowsill in the laboratory where it received ambient sunlight. The following morning, the purple layer had reformed on the side of the tube facing the window at the sediment-water interface. This layer was a physically coherent biofilm with enough structural integrity to withstand the decanting of water and sediment and gentle rinsing with 0.2 μm filtered water from the enrichment site (**Figure 1c**). The biofilm was removed with a sterile toothpick and resuspended with 0.2 μm filtered water from the sampling location. Wet mounts of the resulting concentrated cell suspension revealed a mixed community of organisms dominated by a single 6-10 μm diameter purple coccoid with the characteristic pigmentation of purple sulfur bacteria (**Figure 1d**). Multiple subsequent enrichments from this site, as well as a pool in Little Sippewissett Marsh (Seitz et al., 1993; Shapiro et al., 2011), yielded similar purple biofilms dominated by a morphologically similar organism which we refer to throughout as "*Ca.* T. celerrima".

"*Ca.* T. celerrima" cells were vigorously motile once disturbed from the biofilm (**Movie S1-S3**). Fresh preparations resuspended in 0.2 μm filtered water from the enrichment site were observed to rapidly accumulate in the bottom corner of a well of a 24-well tissue culture plate opposite artificial illumination. This behavior was used as a means to physically enrich "*Ca.* T. celerrima" from associated organisms (**Figure 2a**). Enriched cells contained refractory intracellular globules presumed to be elemental sulfur and a single large lophotrichous flagellar bundle was clearly visible in stationary cells (**Figure 2b**). A small concentrated wet mount of these cells was used to determine whole-cell absorbance spectra using a microscope-mounted spectrophotometer (**Figure 2c**). Spectra of "*Ca.* T. celerrima" best matched those from purple sulfur bacteria of the *Chromatiaceae*, specifically *Chromatium okenii* (Stomp et al., 2007).

Phylogeny and genomic features of "Ca. Thiorhodosphaera celerrima

Aliquots of "*Ca.* T. celerrima" cells purified through at least three rounds of dilution and motility-based concentration from the Ballfield Marsh and Little Sippewisset Marsh sites were pelleted in a microcentrifuge tube and used directly as templates for PCR using universal 16S rRNA gene primers. Sanger sequencing of PCR products yielded two nearly identical 16S sequences. The only variation occurred at two positions where the electrophorograms had ambiguous base calls. The 16S rRNA gene sequence of "*Ca.* T. celerrima" was most closely related to members of the *Chromatiaceae*, but was at most 93.1% identical with a type strain from this family (*Halochromatium salexigens* DSM 4395^{T}). Phylogenetic analysis of 16S rRNA gene sequences from type strains from the *Chromatiaceae* indicate that "*Ca.* T. celerrima" falls within this family (**Figure 3a**). A short 16S rRNA gene fragment was found via a literature search from an enrichment of an unnamed purple sulfur bacteria from Guadeloupe (Gros et al., 2018) to be 99% identical to the 16S of the "*Ca.* T. celerrima". This is the only sequence publicly available with greater than 94% sequence identity to "*Ca.* T. celerrima".

One aliquot of enriched "*Ca.* T. celerrima" cells from Ballfield Marsh was used for whole genome shotgun sequencing. Small amounts of DNA from contaminating organisms was present, so sequencing results were treated as a metagenome and multiple methods of read quality filtering, assembly, and binning were compared using a combination of methods hosted on the KBase platform (Arkin et al., 2018) (see Materials and Methods and **Supplemental file 1** for assembly comparisons). The final "*Ca.* T. celerrima" MAG contained 216 contigs, and was estimated to be 90.5% complete and 1.4% contaminated using the 581 marker gene set for the *Gammaproteobacteria* implemented by CheckM (Parks et al., 2015). This MAG contained a full length 16S rRNA gene sequence matching those produced by

PCR and Sanger sequencing (**Figure 3a**). The PufLM subunits of the photosystem are also used as marker genes for defining the phylogeny of the *Chromatiaceae* (Tank et al., 2009), so we constructed a phylogenetic tree of concatenated PufLM protein sequences and found this too placed "*Ca.* T. celerrima" MAG within the *Chromatiaceae* nearest *Thiorhodovibrio*, albeit with poor branch support. Classification of the "*Ca.* T. celerrima" MAG consistent with the standardized genome taxonomy GTDB (Parks et al., 2018) using GTDB_tk (Chaumeil et al., 2020) placed it within the *Thiorhodovibrio* genus based on its similarity to the single genome representative of this genus *Thiorhodovibrio* sp. 970.

The "*Ca.* T. celerrima" MAG contained features common to the central energy metabolism of purple sulfur bacteria (summarized in **Table S1**). Photosystem genes corresponding to the reaction center (PufLM), tetraheme electron donor for reaction center (PufC), multiple copies of the light harvesting alpha and beta subunits (PufAB), and bacteriochlorophyll biosynthesis genes (BchCXYZ) were present and colocalized in a gene cluster. Cyclic electron transfer between the reaction center occurs through *bc*1 complex (PetABC). Only a *bd*-type terminal oxidase was present (CydAB). Reverse electron transfer is expected to occur for the purpose of generating reduced electron donors NADH, NADPH and Fd. This can be carried out by the Nuo complex (NADH generation from quinols), Rnf (Fd^{2-} generation from NADH), and membrane-bound transhydrogenase (NADPH generation from NADH). These enzymes use proton and sodium motive force to drive these endergonic electron transfer reactions. The genes utilized for sulfur metabolism in the model purple sulfur bacteria *Allochromatium vinosum* (Weissgerber et al., 2014, 2013) are mostly present, including sulfide:quinone oxidoreductase (Sqr), DsrMKJOP membrane complex, sulfite reductase (DsrAB), adenylylsulfate reductase (AprAB), sulfate adenylyltransferase (Sat), and sulfur globule protein (SgpB)(**Table S1**).

Motility of "Ca. Thiorhodosphaera celerrima"

Freshly prepared wet mounts of highly enriched "*Ca.* T. celerrima" cells demonstrated incredibly rapid swimming reminiscent of *Thiovulum majus* (Garcia-Pichel, 1989). To quantify the average swimming speed, high-speed video microscopy analysis was used to track individual cells, recording their path length and time visible (**Figure 4a, Movie S4**). 2,147 individual cells were analyzed in 14 separate acquisitions from two different wet mounts. The average swimming speed was 399 μm/s with a standard deviation 141 μm/s (**Figure 4b**). The fastest cell was observed to transit the entire field of view at an average speed of 814 μm/s. Average cell diameter was found to be 7.49 μm with a standard deviation of 0.76 μm (**Figure 4c**). Swimming speed and cell diameter were not significantly correlated (data not shown).

"*Ca.* T. celerrima" also demonstrated a variety of motility behaviors that have been described previously in members of the purple sulfur bacteria. Concentrated preparations of highly motile cells in 0.2μm filtered source water would rapidly bioconvect (**Figure 5a, Movie S5**). In wells of 24-well plates the cells would accumulate at the air-water interface on the edges of the well, then groups would dislodge from the wall, swimming towards the center of the well before sinking to the bottom. This type of group behavior has been documented in other purple sulfur bacteria such as *C. okenii* (Pfennig, 1968).

Light-based tactic responses were also readily apparent. In wet mounts cells would accumulate at the edge of the water extending beyond the coverslip (**Figure 5b, Movie S6**). When the low intensity microscope light was suddenly increased, the motile cells that had been swarming at the air-water interface immediately dispersed. Decreasing to dim light reversed the behavior and the cells returned to the air-water interface. This behavior could be reproduced repeatedly in a matter of seconds (**Figure**

5b, Movie S6). "*Ca.* T. celerrima" cells were also found to concentrate in locations with steady high levels of illumination compared to their surroundings. This was demonstrated in wet mounts by leaving the transmitted light on through a 10x objective and observing the accumulation of cells in the circle of illumination (**Figure 5c**) and is consistent with the classic "scotophobic" response (Gest, 1995). "*Ca.* T. celerrima" cells often settled and adhered to the surface of whatever container they were prepared in. This was not due to a loss of cell viability as cells were routinely observed to maintain rapid flagellar rotation (**Movie S7**). This adhesion appeared to be a tethering of some kind, as cells were often observed beating their flagella while not being able to break free from a small area (**Movie S8**).

Flow-induced increase in substrate uptake

The rapid motility and biofilm-forming behavior of "*Ca.* T. celerrima" is highly reminiscent of *Thiovulum*. Since these behaviors have been demonstrated to facilitate faster-than-diffusion substrate uptake in *Thiovulum* (Fenchel and Glud, 1998; Petroff et al., 2015), we investigated whether a similar effect could be beneficial in the context of the anoxygenic phototrophic metabolism predicted for "*Ca.* T. celerrima". In the simplest case of free-swimming "*Ca.* T. celerrima" the Reynolds number (*Re*) and the Peclet number (*Pe*) are similar to those for *Thiovulum* (**Table 1**, **Figure 6a**). These dimensionless quantities capture the relative importance of viscous vs. inertial forces and advective vs. diffusive transport, respectively, and illustrate that for both organisms, viscous forces are dominant, and diffusion is the major mechanism of substrate delivery. However, the induced flow is likely to increase substrate uptake as it leads to steeper concentration gradients. At the low *Re* relevant for free-swimming "*Ca.* T. celerrima", the ratio of total to diffusive flux can be estimated as $0.5*[1+(1+2*Pe)^{1/3}]$ (Kiørboe, 2008). Using this approximation for free-swimming "*Ca.* T. celerrima" one would expect a significant increase in substrate delivery by approximately 22%.

The biofilm-forming behavior in the lab and in the environment suggests that the free-swimming calculations above may not accurately reflect the ecophysiological conditions under which this behavior was selected. Assuming that sediment-adhered "*Ca.* T. celerrima" cells use their motility to increase substrate delivery they would need to operate in the opposite direction as *Thiovulum,* pulling up sulfidic water from the sediment instead of pulling down oxygen-rich water, because there is no catabolic role for O_2 in the PSB metabolism.

To capture this scenario we estimate the force a "*Ca.* T. celerrima" cell is able to exert on the surrounding water accounting for the cell size, its velocity and the viscosity of the surrounding fluid (Stoke's law; **Figure 6a**). The resulting force of approximately 31 pN is then used to drive fluid flow in the sediment. We consider flow over 2 cm, the distance between the sediment surface where "*Ca.* T. celerrima" is located and the depth where sulfide profiles reach a steady value of approximately 1mM in the Sippewissett Salt Marsh (Salman et al., 2015). Assuming the force sets up a vertical pressure gradient and a sediment permeability of $10^{-12}m^2$, Darcy's law is used to calculate the flow velocity through the sediment (**Figure 6b**). Using the boundary conditions of [HS-]=0 at the surface, and [HS-]=1 mM at 2 cm depth, we can explicitly solve for the flux of sulfide with and without this pressure-derived flux (see Materials and Methods for details). We find the sulfide delivery to the surface to be increased by ~35% with these assumptions.

While this one-dimensional model captures some basic principles, counter gradients of oxygen and lateral water movement can further complicate the assessment of the impact of flow induced by "*Ca.* T. celerrima" in the environment. For example, biofilms of the close relative of "*Ca.* T. celerrima" from

Guadeloupe revealed an interesting pattern of white sulfur oxidizing bacteria at the margins of the purple mats (Gros et al., 2018). This observation led us to speculate that the upward flux of porewater under the "*Ca.* T. celerrima" mat might improve the energetics of colorless sulfide oxidizers at the margins of "*Ca.* T cellerima" biofilms due to the recharge with oxygen-rich surface water at the biofilm margins. To assess this, we implemented a radially symmetric advection-diffusion-reaction model comprised of a circular "*Ca.* T. cellerima" biofilm surrounded by a sulfide oxidizing biofilm, with 1mM sulfidic conditions imposed 2 cm deep in the sediment and oxic conditions in the overlying water (Materials and Methods; **Figure 6c**). By comparing the sulfide consumption in the central region of the "*Ca.* T. cellerima" biofilm with and without flow (**Figure 6d** and **e**), we find that flow increases sulfide consumption by 37%. Contrary to our expectations, the model shows little effect of sulfide or O_2 consumption rates in the surrounding chemotrophic mat.

Discussion

"Ca. Thiorhodosphaera celerrima" represents a novel genus within the Chromatiaceae

16S rRNA gene sequence identity values of 94.5 -95% have been proposed for delineating genus-level divisions within bacteria (Tindall et al., 2010; Yarza et al., 2014). It is also suggested that "these genetic criteria should always be accompanied by a discriminant phenotypic property" (Yarza et al., 2014). By both metrics "*Ca.* T. celerrima" should be assigned to a new genus. All phytogenetic analyses clearly place these organisms within the *Chromatiaceae* family (**Figure 3**) but 16S rRNA sequence identity values to any currently recognized members are ≤93%. Phenotypically they can be easily distinguished from other genera within the *Chromatiaceae* as their cell diameter of ~7.5 μm is larger than currently described

genera within the family (Imhoff, 2014) and their swimming speeds exceed anything known for a photosynthetic organism, a feature which doubles as a useful method for cell enrichment from complex communities. Based on a short 16S rRNA gene sequence it was suggested that a close relative of "*Ca.* T. celerrima" from Guadeloupe be assigned to the genus *Halochromatium* (Gros et al., 2018). With the expanded phylogenetic analysis provided here, we suggest both of these organisms are species within the new genera "*Ca.* Thiorhodosphaera". Characteristics of both organisms are consistent including cell size, lophotrichous flagellar bundles, and photophobic behavior (Gros et al., 2018). The Guadeloupe strain was also characterized ultrastructurally and the intracellular sulfur globules were verified using EDS (Gros et al., 2018). Taxonomic characterization based on the GTDB_tk algorithm assigns "*Ca.* T. celerrima" to the *Thiorhodovibrio* genus based on similarity to the *Thiorhodovibrio* sp. 970 genome. While we recognize the value of such methods for classifying MAGs with no discriminant phenotypic property, we find the differences apparent between "*Ca.* Thiorhodosphaera" described here and previously (Gros et al., 2018) to be reason enough to propose a new genus separate from the *Thiorhodovibrio* (Imhoff and Overmann, 2015).

"*Candidatus* Thiorhodosphaera celerrima". Gr. n. theion (Latin transliteration thium), sulfur; Gr. n. rhodon, the rose; L. fem. n. sphaera, a ball, sphere; N.L. fem. n. Thiorhodosphaera, the red sulfur ball. L. fem. adj. celerrima, superlative degree of celer, fast. "*Candidatus* Thiorhodosphaera celerrima", the fastest red sulfur ball.

Ecological implications of fast swimming in "Ca. Thiorhodosphaera celerrima"

Our phylogenetic and genomic analyses, spectroscopic characterization, and previous ultrastructural and elemental analysis (Gros et al., 2018), suggest "*Ca.* T. celerrima" is a sulfur oxidizing phototroph,

although this has not been shown conclusively by growth under these conditions. Culturing attempts were undertaken starting with enrichments of "*Ca.* T. celerrima" in media designed for the isolation of sulfide oxidizing *Chromatiaceae* with an infrared LED light source, but we were unsuccessful at producing a pure culture. The closely related organism from Guadeloupe also resisted cultivation (Gros et al., 2018).

Both colorless sulfur bacteria and purple sulfur bacteria use reduced sulfur compounds as electron donors, but only the former are commonly found to have morphological and behavioral adaptations to improve their substrate consumption beyond the limitations of diffusion (Fenchel and Glud, 1998; Huettel et al., 1996; Schulz and Jørgensen, 2001). This difference could be rationalized on metabolic grounds. The purple sulfur bacteria employ cyclic electron flow, which uses light energy to run a single electron many times through their photosystem and electron transport chain. New electrons are only needed to replace those lost to anabolic reactions such as carbon or nitrogen fixation. Each reduced sulfur compound that is oxidized therefore produces a large non-stoichiometric amount of ATP. In colorless sulfur bacteria each electron is transferred to the terminal electron acceptor in a linear manner producing a roughly stoichiometric amount of ATP per electron donor and acceptor. This difference between the amount of energy gained per substrate molecule may explain why it has been more common to find colorless sulfur bacteria with adaptations aimed at increasing their substrate flux.

While not stoichiometrically dependent on sulfide for ATP, purple sulfur bacteria are still intimately tied to the supply of reduced sulfur compounds. They can differentially respond to diurnal cycles of sulfide profiles and light levels (van Gemerden, 1974), with studies in marine sulfuretums demonstrating that some members of the *Chromatiaceae* can migrate between the sediment and water to follow optimal concentration gradients of sulfide, oxygen and light (Jørgensen et al., 1982). Laboratory competition

studies show that slight differences between the kinetics of sulfide uptake or the ability to store intermediate sulfur compounds for later use are important niche-defining characteristics (Van Gemerden and Mas, 1995). While electron donor limitation may not be as energetically costly to purple sulfur bacteria as it is in the colorless sulfur bacteria, considerations surrounding acquisition of electron donor seem central to the adaptations that define the different purple sulfur bacteria niches.

Based on our analysis of flow-induced sulfide delivery to "*Ca.* T. celerrima" we could potentially expect uptake to increase by 30-40%, depending on the assumptions and models being used. While these values are very approximate, in principle it seems that the same adaptive advantage explaining *Thiovulum*'s extreme motility could apply to "*Ca.* T. celerrima" as well. Normal purple sulfur bacteria face a trade-off when choosing where to position themselves in sediment or water columns. Closer to the surface will provide more light energy but less electron donor, while deeper will yield greater electron donor but less light energy to oxidize it with. The rapid motility demonstrated here may allow "*Ca.* T. celerrima" to get the best of both options, high light energy and increased sulfide supply, with some of the increased energy production being used to drive fluid flow.

The top speed of 814 μm/s makes "*Ca.* T. celerrima" the second fastest bacterium yet described, only exceeded by "*Candidatus* Ovobacter propellens" found to have a maximum speed of approximately 1mm/s. While no information on the physiology or phylogeny of "*Ca.* O. propellens" is available, it was found associated with *Thiovulum*-containing veils in sulfidic sediments (Fenchel and Thar, 2004). "*Ca.* T. celerrima" from Woods Hole, MA and Guadeloupe were both found in habitats that are routinely found to contain *Thiovulum* veils (Garcia-Pichel, 1989, p.; Gros, 2017; Gros et al., 2018; Petroff et al., 2015). Clearly, there appears to be a strong selective advantage to this rapid motility phenotype in these sulfidic

environments, suggesting this exceptional swimming speed may be an important behavioral adaptation to the gradient lifestyle. While "*Ca.* T. celerrima" remains to be cultured, we hope that the generation of a draft genome and phylogenetic markers linked to morphologic description and a simple enrichment strategy will facilitate more studies of this interesting new phototroph. Whether the swimming speed is used in the manner hypothesized here remains to be determined, and such ideas could be tested by measuring the flow fields above dense biofilms using micron-sized latex beads, as done previously for *Thiovulum* (Fenchel and Glud, 1998).

Materials and Methods

Sampling locations

Samples of "*Ca.* T. celerrima" were collected from the Ballfield Marsh behind Woods Hole park (41.529722 N 70.670242 W) and Little Sippewissett Marsh (41.576373 N 70.638566 W) in July, 2016. Samples of surface sediment and overlying water were collected in 50ml polypropylene tubes. An additional ~1L of the overlying water was collected and sterile filtered (0.2μm) to be used for serial enrichment and wet mounts.

Cell enrichment

The general strategy for cell enrichment is illustrated in **Figure 2a**. Free-swimming cells of "*Ca.* T. celerrima" were prepared by gently resuspending biofilms formed on the side of polypropylene containers in 0.2μm filtered source water in a well of a 24-well plate (**Figure 1c**). These cells were found to rapidly swim away from intense artificial light, forming a dense pellet in the opposing corner of the well. At least four wells of a 24-well plate were filled with additional sterile-filtered source water and the concentrated "*Ca.* T. celerrima" population was pipetted into new sterile well. Enrichments were

sequentially resuspend and concentrated in this way until no other cell morphologies were observed (typically 3-4 passages).

Spectroscopy

A concentrated preparation of "*Ca.* T. celerrima" was prepared on a microscope slide and absorbance measurements were made with a fiber optic spectrophotometer attached to a microscope eye piece. Background light spectrum of the lamp without cells was substracted from the spectrum of light passing through "*Ca.* T. celerrima".

16S rRNA gene sequencing

A preparation of "*Ca.* T. celerrima" from Ballfield Marsh and Woods Hole were pelleted in a microcentrifuge and used as templates for a colony PCR using the 8F and 1391R universal 16S primer set (8F: agagtttgatcctggctcag 1391R: gacgggcggtgwgtrca), with an extended initial step to facilitate cell lysis: 10 minutes at 98°C; 30 cycles of: 10s at 98°C; 30s at 60°C; 45s at 72°C; and 5 minutes of 72°C. PCR products were gel purified and sequenced using sanger sequencing. Sequences were deposited on GenBank with accession numbers MN559830.1 and MN559831.1.

DNA Extraction for whole genome sequencing

A frozen pellet of "*Ca.* T. celerrima" from Ballfield Marsh from the same enriched preparation that was analyzed for motility was used for DNA isolation using the following protocol. Pellet was resuspended in 100ul of lysis buffer (20mM Tris, 1mM EDTA, 100mM NaCl, 15mg/ml lysozyme, 1mg/ml RNAse). Resuspended cells were heated at 37°C for 30 minutes, then 10ul of 10% SDS was added and the solution vortexed for 1 minute, followed by a standard phenol:chloroform:isoamyl alcohol extraction.

The aqueous phase was combined with 1/10th volume of 3M sodium acetate (pH 5.2), and 2.5x volumes of 100% ethanol and 10ul of linear acrylimide (10mg/ml). The resulting solution was precipitated -20°C for an hour and then spun at max speed in a tabletop centrifuge at 4°C for 30 minutes. The resulting pellet was washed with chilled 70% ethanol and pelleted at max speed in a tabletop centrifuge at 4°C for 5 minutes. Ethanol was removed with a pipette and tube cap left open for residual to evaporate. DNA was resuspended in nuclease free water. DNA was sequenced on an Illumina HiSeq 2500 with NEBNext® Multiplex library with 700bp insert size. Sequencing was carried out by the Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech and produced ~19 million reads of 250bp paired-end sequences.

Genome assembly

Sequence data was processed into four datasets with different quality filtering including: raw data, Illumina chastity filtering, and Trimmomatic applied with two different settings (see **Supplemental file 1** for details). These four datasets were assembled with four different assemblers available on the Kbase platform (Arkin et al., 2018): SPAdes (Bankevich et al., 2012), metaSPAdes (Nurk et al., 2017), Kiki (https://github.com/GeneAssembly/kiki) and MEGAHIT (Li et al., 2015). All 16 resulting assemblies were then binned with metaBAT2 (Kang et al., 2019). This process only resulted in a single dominant bin in all cases, which was quality assessed using checkM (Parks et al., 2015). The estimated completeness and contamination for all assemblies is reported in **Supplemental file 1**. The assemblies were routinely best with Kiki based on high completeness, low contamination, and length of contigs. The best assembly was selected and manually binned using the anvi'o workflow for metagenome binning (with only a single sample for coverage information) (Eren et al., 2015). This resulted in a final MAG of 3.6 million bp on

216 contigs with 90.5% completeness and 1.4% contamination. This MAG was annotated with the IMG webserver (Chen et al., 2019) and can be found with Taxon Object ID: 3300034821.

Phylogentic reconstruction

16S rRNA and PufLM phylogenies have been used to define phylogenetic relationships within the family Chromatiaceae in references (Imhoff, 2014) and (Tank et al., 2009), respectively. 16S rRNA genes from type species of the Chromatiaceae and "*Ca.* T. celerrima" were aligned with the SINA aligner (Pruesse et al., 2012) with SILVA database 132 (Quast et al., 2013). 16S tree was built on the webserver http://www.atgc-montpellier.fr/phyml/ running PhyML 3.0 (Guindon et al., 2010). Substitution model was automatically selected using Akaike Information Criterion implemented by SMS (Lefort et al., 2017) and the tree was constructed with 100 bootstrap replicates for branch support. Concatenated PufLM sequences were aligned using muscle 3.8.31 (Edgar, 2004) and three was constructed using RAxML 8.2.12 (Stamatakis, 2014) implemented on the CIPRES Science Gateway (Miller et al., 2010). The CAT evolutionary model (Si Quang et al., 2008) was used with 25 distinct rate categories and DAYHOFF substitution matrix. 100 bootstrap replicates were calculated for branch support.

High speed video microscopy image analysis

At least four rounds of motility-based enrichment was carried out on a sample of "*Ca.* T. celerrima" from Ballfield Marsh before microscopy analysis. Wet mounts of enriched cells were visualized with a ZEISS Axio Observer Z1 with two different objectives: Plan-Neofluar 10x/0.30 Ph1 and Plan-Apochromat 20x/0.8 M27. Video captures were taken at 120 frames per second in three second bursts with ZEISS Axiocam 702 mono. After analysis pixel/micron calibration was made for both objectives by capturing images of a micrometer.

Image stacks originally captured in .czi file format were exported as individual TIFF images. TIFF images were imported into ImageJ as an image stack and tracks were determined using TrackMate v3.4.2 (Tinevez et al., 2017). The only image manipulation required before applying TrackMate was inversion so cells appeared as increased pixel intensity on a dark background. LoG (Laplacian of Gaussian) particle detection was used with 8μm diameter and threshold value of 2. Tracks were determined using the simple LAP (Linear assignment problem) trackers. Raw track data was exported as xml file and further processed in MATLAB 2016 (MathWorks, Natick, MA, USA). Some cells were either adhered to the microscope slide or present in only a few frames of the analysis so a 10 body length displacement (100μm) minimum cutoff was applied to tracks.

Flow modeling

Values of Reynolds number (*Re*) and Peclet number (*Pe*) in **Table 1** were calculated using $Re = u*r/v$ and $Pe = u*r/D$ where r = cell radius, u = swimming speed, v = viscosity (~10^-6 m^2/s) and D = diffusion coefficient (~10^-9 m^2/s) (Kiørboe, 2008). The ratio of substrate delivery with flow vs. stationary was calculated for free swimming cells from the Sherwood number approximation (*Sh*): $Sh = 0.5*[1+(1+2*Pe)^{1/3}]$.

The force free-swimming "*Ca.* T. celerrima" exert is calculated from Stoke's law ($F = 6\pi\mu ru$) with μ = dynamic viscosity ($1.09*10^{-3}$ Pa*s) and r and u as above, yielding ~31pN (similar to the 40pN found for *Thiovulum* in reference (Petroff et al., 2015)). Darcy's law ($q = -k/\mu * \Delta p/L$) is used to calculate the fluid flow rate (q) based on a pressure gradient (Δp), permeability ($k = 10^{-12} m^2$) and distance (L=0.02m). The pressure gradient is calculated assuming pressure = 0 at 2cm depth and pressure at the surface from the

force (31pN) and cross-sectional area of a "*Ca.* T. celerrima" cell. See **Supplementary file 2** for full calculations.

The 1D model depicted in **Figure 6b** considers a 2cm vertical sediment column with constant porosity and no sources or sinks of sulfide between the top and bottom and fluid flow (u) calculated as above. The sulfide concentrations are set as boundary conditions C(x=0) = C_{top} = 0 and C(x=2cm) = C_{bot} = 1mM. The convection-diffusion equation at steady state (dC/dt=0) is given in equation (eq. 1)

$$\frac{dC}{dt} = 0 = D\frac{d^2C}{dx^2} - u\frac{dC}{dx} \quad (1)$$

Which can be solved explicitly for sulfide concentration throughout the sediment column (eq. 2). The flux through the sediment profile due to diffusion and advection are shown in (eq. 3 and 4).

$$C(x) = \frac{C_{top}e^{u\frac{L}{D}} + (C_{bot} - C_{top})e^{u\frac{x}{D}} - C_{bot}}{e^{u\frac{L}{D}} - 1} \quad (2)$$

$$F_{diff} = \frac{-(C_{bot} - C_{top})ue^{u\frac{x}{D}}}{e^{u\frac{L}{D}} - 1} \quad (3)$$

$$F_{adv} = u\left(-\frac{C_{bot} - C_{top}e^{u\frac{L}{D}}}{e^{u\frac{L}{D}} - 1}\right) \quad (4)$$

With these conditions, the flux based on diffusion at the surface is ~6.76 mol/m^2/s, and this diffusive flux is much greater than the advective flux near the surface (with C_{top} = 0 the advective flux at the surface is zero). The dominance of diffusion is expected based on the low *Re* and *Pe* values. The diffusive flux in this gradient without flow is found with Fick's law (J = -DdC/dx) with D given above, dC = 1mM and dx = 2cm to be 5mol/m^2/s. The ratio of the diffusive flux in the with flow to diffusive flux without flow is presented in Table 1 is therefore 6.76/5 = 1.35.

The 3D model geometry was motivated by the images of a close relative of "*Ca.* T. celerrima" from Guadeloupe where a mat of purple sulfur bacteria was surrounded by a ring of while sulfide oxidizing bacteria (Gros et al., 2018). The radially symmetric advection-diffusion-reaction model was implemented in COMSOL Multiphysics 5.4 (COMSOL Inc., Burlington, MA, USA). Details of the modeling assumptions, domain properties, and modelling output can be found in **Supplementary files 3 and 4**. The consumption of sulfide in domain 1 corresponding to a "*Ca.* T. celerrima" biofilm was compared with and without flow for the values presented in **Table 1**.

Acknowledgements

This research was performed in conjunction with the Woods Hole microbial diversity course of 2016. We would like to thank Megan Bergkessel for assistance with the DNA extraction, Kate Hargreaves for 16S sequencing, and Igor Antoshechkin of the Millard and Muriel Jacobs Genetics and Genomics Laboratory at Caltech for library preparation and sequencing. G.L.C. was supported by the NIH/NRSA training grant, T32 GM007616.

Conflict of Interest

The authors declare no conflict of interest

References

Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald

PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology* **36**:566–569. doi:10.1038/nbt.4163

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**:455–477.

Chaumeil P-A, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **36**:1925–1927. doi:10.1093/bioinformatics/btz848

Chen I-MA, Chu K, Palaniappan K, Pillay M, Ratner A, Huang J, Huntemann M, Varghese N, White JR, Seshadri R, Smirnova T, Kirton E, Jungbluth SP, Woyke T, Eloe-Fadrosh EA, Ivanova NN, Kyrpides NC. 2019. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Research* **47**:D666–D677. doi:10.1093/nar/gky901

Decker KLM, Potter CS, Bebout BM, Des Marais DJ, Carpenter S, Discipulo M, Hoehler TM, Miller SR, Thamdrup B, Turk KA, Visscher PT. 2005. Mathematical simulation of the diel O, S, and C biogeochemistry of a hypersaline microbial mat. *FEMS Microbiol Ecol* **52**:377–395. doi:10.1016/j.femsec.2004.12.005

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–1797. doi:10.1093/nar/gkh340

Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**:e1319. doi:10.7717/peerj.1319

Fenchel T, Glud RN. 1998. Veil architecture in a sulphide-oxidizing bacterium enhances countercurrent flux. *Nature* **394**:367–369. doi:10.1038/28609

Fenchel T, Thar R. 2004. "Candidatus Ovobacter propellens": a large conspicuous prokaryote with an unusual motility behaviour. *FEMS Microbiology Ecology* **48**:231–238. doi:10.1016/j.femsec.2004.01.013

Fossing H, Gallardo VA, Jørgensen BB, Hüttel M, Nielsen LP, Schulz H, Canfield DE, Forster S, Glud RN, Gundersen JK, Küver J, Ramsing NB, Teske A, Thamdrup B, Ulloa O. 1995. Concentration and transport of nitrate by the mat-forming sulphur bacterium Thioploca. *Nature* **374**:713–715. doi:10.1038/374713a0

Garcia-Pichel F. 1989. Rapid bacterial swimming measured in swarming cells of Thiovulum majus. *J Bacteriol* **171**:3560–3563. doi:10.1128/jb.171.6.3560-3563.1989

Gest H. 1995. Phototaxis and other sensory phenomena in purple photosynthetic bacteria. *FEMS Microbiol Rev* **16**:287–294. doi:10.1111/j.1574-6976.1995.tb00176.x

Gros O. 2017. First description of a new uncultured epsilon sulfur bacterium colonizing marine mangrove sediment in the Caribbean: Thiovulum sp. strain karukerense. *FEMS Microbiol Lett* **364**. doi:10.1093/femsle/fnx172

Gros O, Bisqué L, Sadjan M, Azede C, Jean-Louis P, Guidi-Rontani C. 2018. First description of a new uncultured purple sulfur bacterium colonizing marine mangrove sediment in the Caribbean: Halochromatium-like PSB from Guadeloupe. *Comptes Rendus Biologies* **341**:387–397. doi:10.1016/j.crvi.2018.07.001

Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Syst Biol* **59**:307–321. doi:10.1093/sysbio/syq010

Huettel M, Forster S, Kloser S, Fossing H. 1996. Vertical Migration in the Sediment-Dwelling Sulfur Bacteria Thioploca spp. in Overcoming Diffusion Limitations. *Appl Environ Microbiol* **62**:1863–1872.

Imhoff JF. 2014. The Family Chromatiaceae In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes: Gammaproteobacteria. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 151–178. doi:10.1007/978-3-642-38922-1_295

Imhoff JF, Overmann J. 2015. Thiorhodovibrio. *Bergey's Manual of Systematics of Archaea and Bacteria* 1–6.

Jørgensen BB, Des Marais DJ. 1986. Competition for sulfide among colorless and purple sulfur bacteria in cyanobacterial mats. *FEMS Microbiol Ecol* **2**:179–186. doi:10.1111/j.1574-6968.1986.tb01727.x

Jørgensen BB, Postgate JR, Postgate JR, Kelly DP. 1982. Ecology of the bacteria of the sulphur cycle with special reference to anoxic—oxic interface environments. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* **298**:543–561. doi:10.1098/rstb.1982.0096

Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **7**:e7359. doi:10.7717/peerj.7359

Kiørboe T. 2008. A mechanistic approach to plankton ecology. Princeton University Press.

Lefort V, Longueville J-E, Gascuel O. 2017. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* **34**:2422–2424. doi:10.1093/molbev/msx149

Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**:1674–1676.

Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees2010 Gateway Computing Environments Workshop (GCE). Presented at the 2010 Gateway Computing Environments Workshop (GCE). pp. 1–8. doi:10.1109/GCE.2010.5676129

Møller MM, Nielsen LP, Jørgensen BB. 1985. Oxygen Responses and Mat Formation by Beggiatoa spp. *Appl Environ Microbiol* **50**:373–382.

Murray RGE, Stackebrandt E. 1995. Taxonomic Note: Implementation of the Provisional Status Candidatus for Incompletely Described Procaryotes. *International Journal of Systematic and Evolutionary Microbiology,* **45**:186–187. doi:10.1099/00207713-45-1-186

Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res* **27**:824–834.

Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* **36**:996–1004. doi:10.1038/nbt.4229

Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**:1043–1055.

Petroff AP, Pasulka AL, Soplop N, Wu X-L, Libchaber A. 2015. Biophysical basis for convergent evolution of two veil-forming microbes. *Royal Society Open Science* **2**:150437. doi:10.1098/rsos.150437

Pfennig N. 1968. Chromatium okenii (Thiorhodaceae) Biokonvektion, aero- und phototaktisches Verhalten. *ENCYCLOPAEDIA CINEMATOGRAPHICA*.

Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**:1823–1829.

Purcell EM. 1977. Life at low Reynolds number. *American Journal of Physics* **45**:3–11. doi:10.1119/1.10903

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Gl$\surd$$\partial$ckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**:D590–D596.

Salman V, Yang T, Berben T, Klein F, Angert E, Teske A. 2015. Calcite-accumulating large sulfur bacteria of the genus Achromatium in Sippewissett Salt Marsh. *The ISME Journal* **9**:2503–2514. doi:10.1038/ismej.2015.62

Schulz HN, Beer D de. 2002. Uptake Rates of Oxygen and Sulfide Measured with Individual Thiomargarita namibiensis Cells by Using Microelectrodes. *Appl Environ Microbiol* **68**:5746–5749. doi:10.1128/AEM.68.11.5746-5749.2002

Schulz HN, Jørgensen BB. 2001. Big Bacteria. *Annu Rev Microbiol* **55**:105–137.

Seitz AP, Nielsen TH, Overmann J. 1993. Physiology of purple sulfur bacteria forming macroscopic aggregates in Great Sippewissett Salt Marsh, Massachusetts. *FEMS Microbiology Ecology* **12**:225–235. doi:10.1111/j.1574-6941.1993.tb00035.x

Shapiro OH, Hatzenpichler R, Buckley DH, Zinder SH, Orphan VJ. 2011. Multicellular photo-magnetotactic bacteria. *Environmental Microbiology Reports* **3**:233–238. doi:10.1111/j.1758-2229.2010.00215.x

Si Quang L, Gascuel O, Lartillot N. 2008. Empirical profile mixture models for phylogenetic reconstruction. *Bioinformatics* **24**:2317–2323. doi:10.1093/bioinformatics/btn445

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313.

Stomp M, Huisman J, Stal LJ, Matthijs HCP. 2007. Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. *ISME J* **1**:271–282. doi:10.1038/ismej.2007.59

Tank M, Thiel V, Imhoff JF. 2009. Phylogenetic relationship of phototrophic purple sulfur bacteria according to pufL and pufM genes. *International Microbiology* **12**:175–185.

Thar R, Kühl M. 2001. Motility of Marichromatium gracilein Response to Light, Oxygen, and Sulfide. *Appl Environ Microbiol* **67**:5410–5419. doi:10.1128/AEM.67.12.5410-5419.2001

Tindall BJ, Rosselló-Móra R, Busse H-J, Ludwig W, Kämpfer P. 2010. Notes on the characterization of prokaryote strains for taxonomic purposes. *International Journal of Systematic and Evolutionary Microbiology,* **60**:249–266. doi:10.1099/ijs.0.016949-0

Tinevez J-Y, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, Bednarek SY, Shorte SL, Eliceiri KW. 2017. TrackMate: An open and extensible platform for single-particle tracking. *Methods*, Image Processing for Biologists **115**:80–90. doi:10.1016/j.ymeth.2016.09.016

van Gemerden H. 1974. Coexistence of organisms competing for the same substrate: An example among the purple sulfur bacteria. *Microb Ecol* **1**:104–119. doi:10.1007/BF02512382

van Gemerden H, Beeftink HH. 1981. Coexistence of Chlorobium and Chromatium in a sulfide-limited continuous culture. *Arch Microbiol* **129**:32–34. doi:10.1007/BF00417175

Van Gemerden H, Mas J. 1995. Ecology of Phototrophic Sulfur Bacteria In: Blankenship RE, Madigan MT, Bauer CE, editors. Anoxygenic Photosynthetic Bacteria, Advances in Photosynthesis and Respiration. Dordrecht: Springer Netherlands. pp. 49–85. doi:10.1007/0-306-47954-0_4

Visscher PT, van den Ende FP, Schaub BEM, van Gemerden H. 1992. Competition between anoxygenic phototrophic bacteria and colorless sulfur bacteria in a microbial mat. *FEMS Microbiol Lett* **101**:51–58. doi:10.1111/j.1574-6968.1992.tb05761.x

Weissgerber T, Dobler N, Polen T, Latus J, Stockdreher Y, Dahl C. 2013. Genome-Wide Transcriptional Profiling of the Purple Sulfur Bacterium Allochromatium vinosum DSM 180T during Growth on Different Reduced Sulfur Compounds. *Journal of Bacteriology* **195**:4231–4245. doi:10.1128/JB.00154-13

Weissgerber T, Sylvester M, Kröninger L, Dahl C. 2014. A Comparative Quantitative Proteomic Study Identifies New Proteins Relevant for Sulfur Oxidation in the Purple Sulfur Bacterium Allochromatium vinosum. *Appl Environ Microbiol* **80**:2279–2292. doi:10.1128/AEM.04182-13

Wit R de, Ende FP van den, Gemerden H van. 1995. Mathematical simulation of the interactions among cyanobacteria, purple sulfur bacteria and chemotrophic sulfur bacteria in microbial mat communities. *FEMS Microbiology Ecology* **17**:117–136. doi:10.1111/j.1574-6941.1995.tb00136.x

Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, Whitman WB, Euzéby J, Amann R, Rosselló-Móra R. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology* **12**:635–645. doi:10.1038/nrmicro3330

Figures

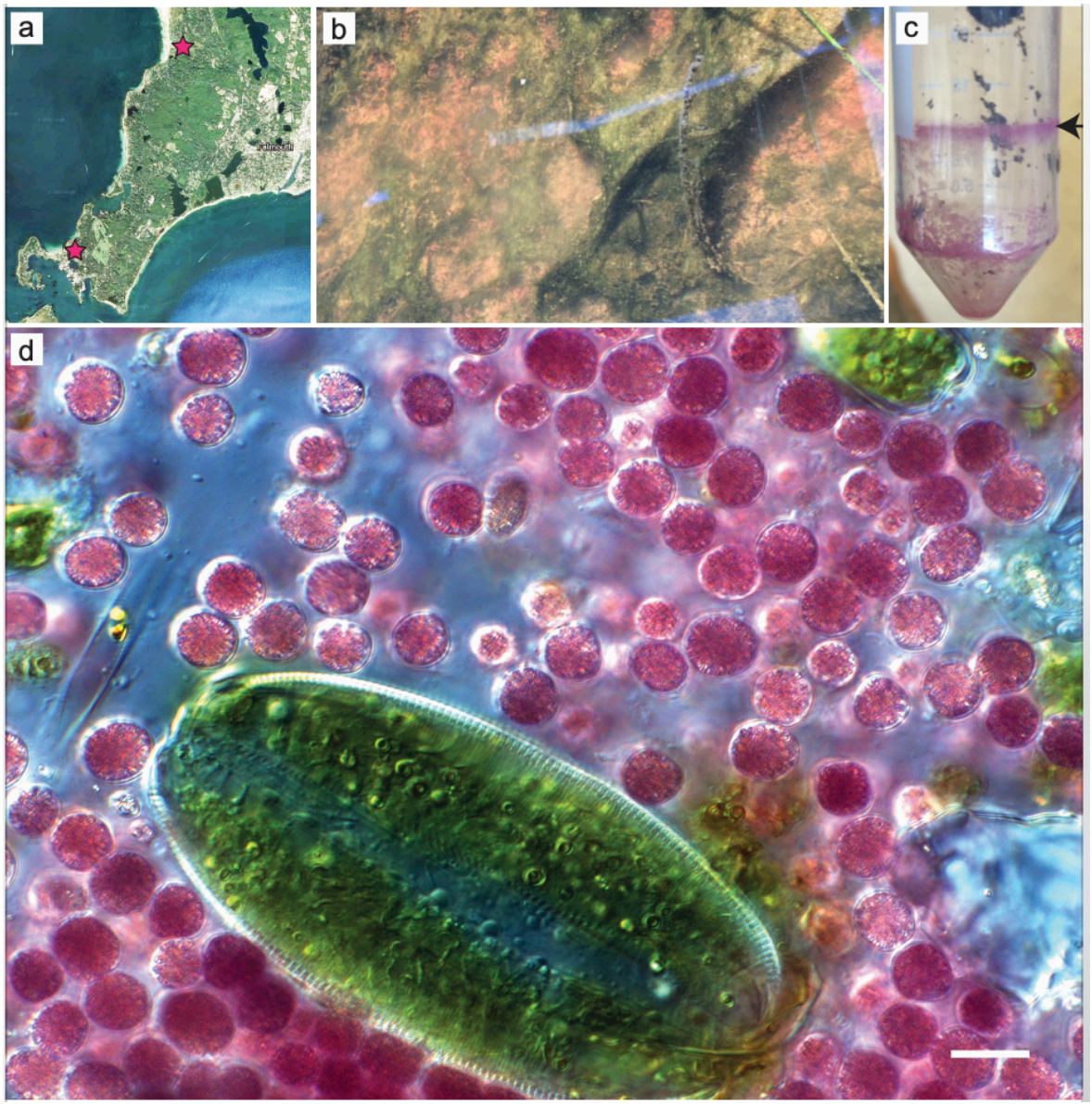


Figure 1 Habitat and initial observations of "*Ca.* T. celerrima" (**A**) Satellite imagery of Woods Hole, Ma. and nearby areas obtained from Google Earth. Stars indicate sampling locations, bottom left marks Ballfield Marsh behind Woods Hole park, and top right marks Little Sippewissett Marsh. (**B**) Characteristic image of purple layer overlying sediments found in Ballfield Marsh. (**C**) After one night in the laboratory, purple layer moved and adhered to the side of the sealed polypropylene tube. The biofilm was strongly attached to the wall of the tube allowing for easy decanting of sediment and source water. Arrow indicates the top level of sediment where the majority of the cells would accumulate in these enrichments. (**D**) Wet mount produced from biofilm community scraped from the side of the polypropylene tube reveal a single dominant purple morphotype with many other associated photosynthetic organisms. Scale bar 10 μm.

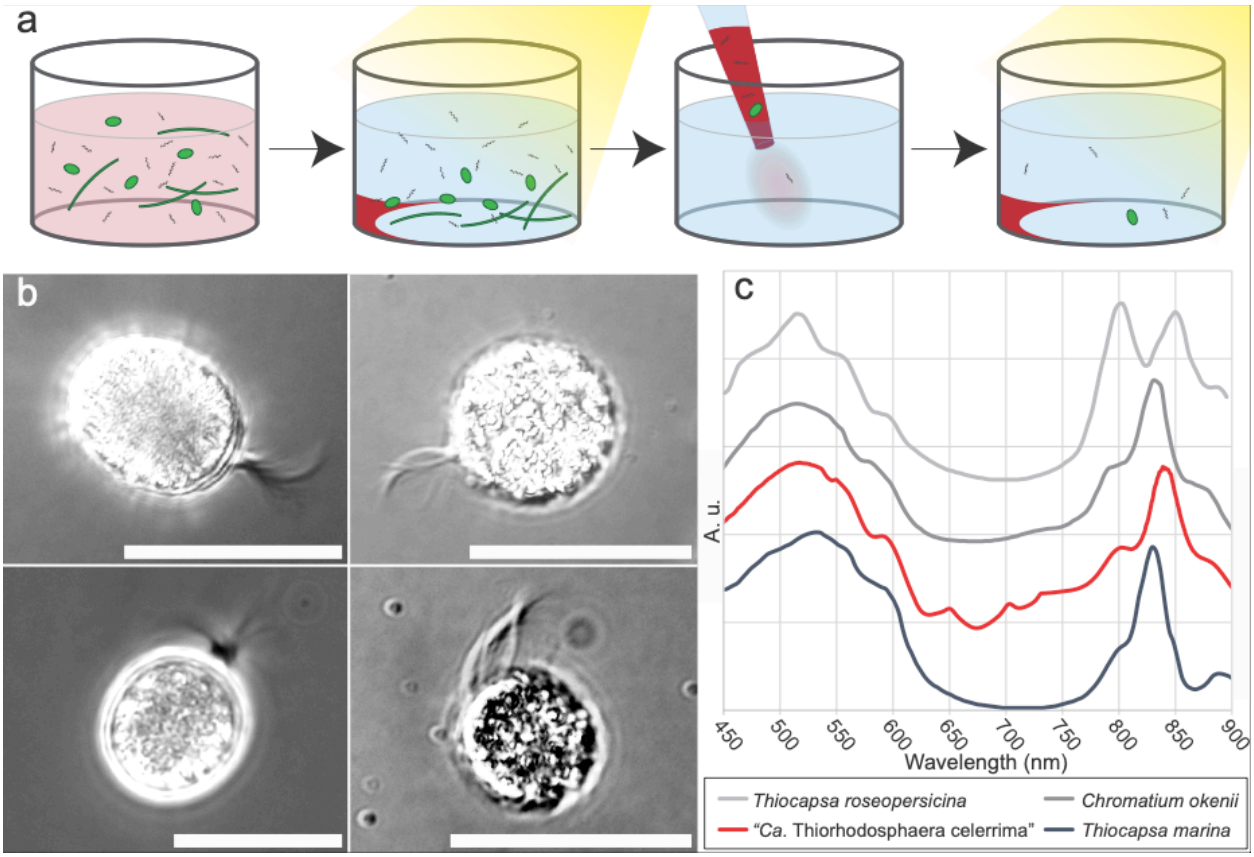


Figure 2 Enrichment and physical characteristics of "*Ca.* T. celerrima". (**A**) Extremely rapid light-based motility allowed for the preparation of highly purified enrichments of "*Ca.* T. celerrima". Biofilms adhered to sampling containers as in **Figure 1C** were scraped and resuspended in filtered source water resulting in a mixed enrichment. Successive rounds of illumination and pipetting of self-concentrated "*Ca.* T. celerrima" cells into new sterile media progressively increased purity. (**B**) Characteristic photomicrographs of "*Ca.* T. celerrima" cells demonstrating lophotrichous flagellar tufts and refractile sulfur globules. Scale bars all 10 μm. **c** Absorbance spectra of a dense preparation of "*Ca.* T. celerrima" cells with characteristic whole cell spectra of related purple sulfur bacteria from ref. 21.

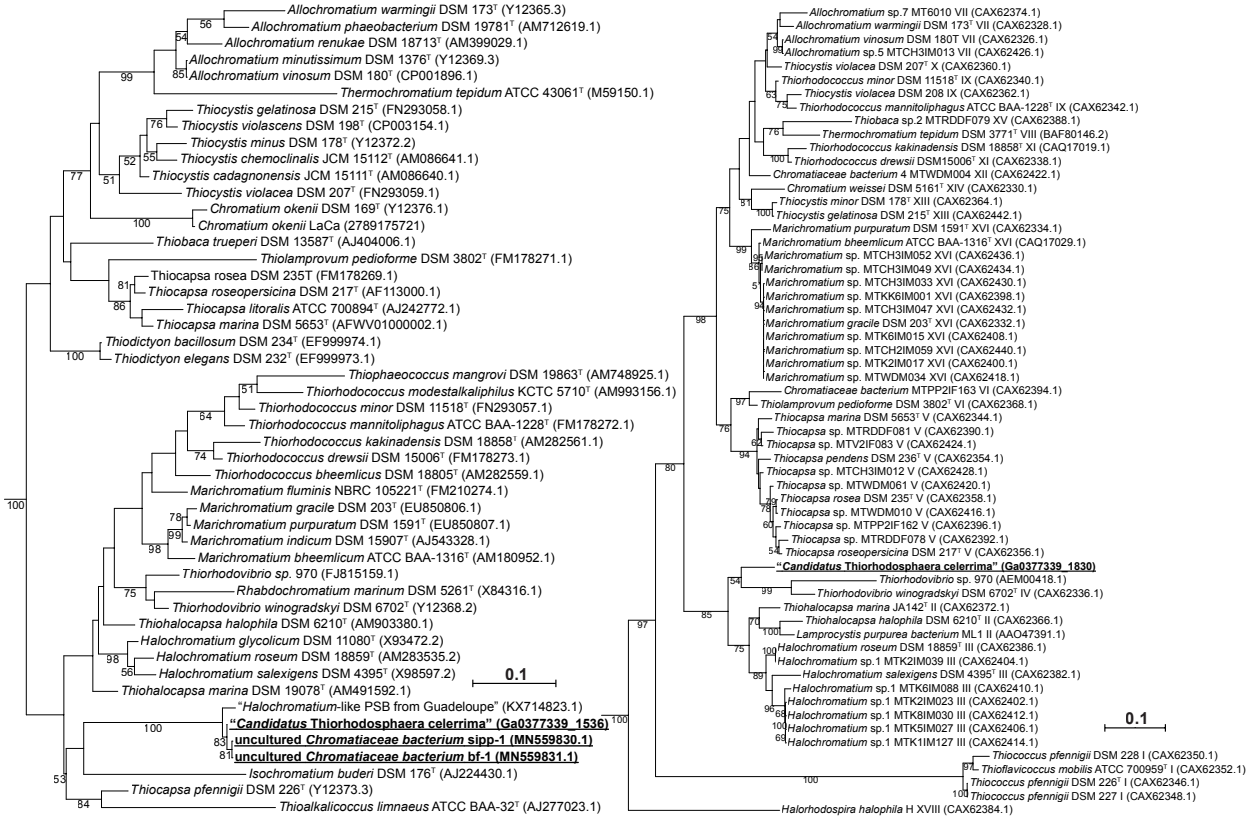


Figure 3 Phylogenetic analysis of "*Ca.* T. celerrima" and related organisms. (**A**) Phylogenetic tree constructed from 16S rRNA gene sequences of type species from all major genera within the family *Chromatiaceae*. Alignments were generated with the SINA Aligner using the global SILVA alignment and tree was calculated using PhyML (see **Materials and Methods** for details). (**B**) Phylogenetic tree constructed from concatenated protein sequences of the photosynthetic reaction center PufLM also places "*Ca.* T. celerrima" within the *Chromatiaceae*. Accession numbers are GenBank IDs except for "*Ca.* T. celerrima" and *Chromatium okenii* LaCa which are IMG IDs. Tree scales represents substitutions per site, and bootstrap branch support values above 50 are shown. Roots leads to members of the related *Ectothiorhodospiraceae* family.

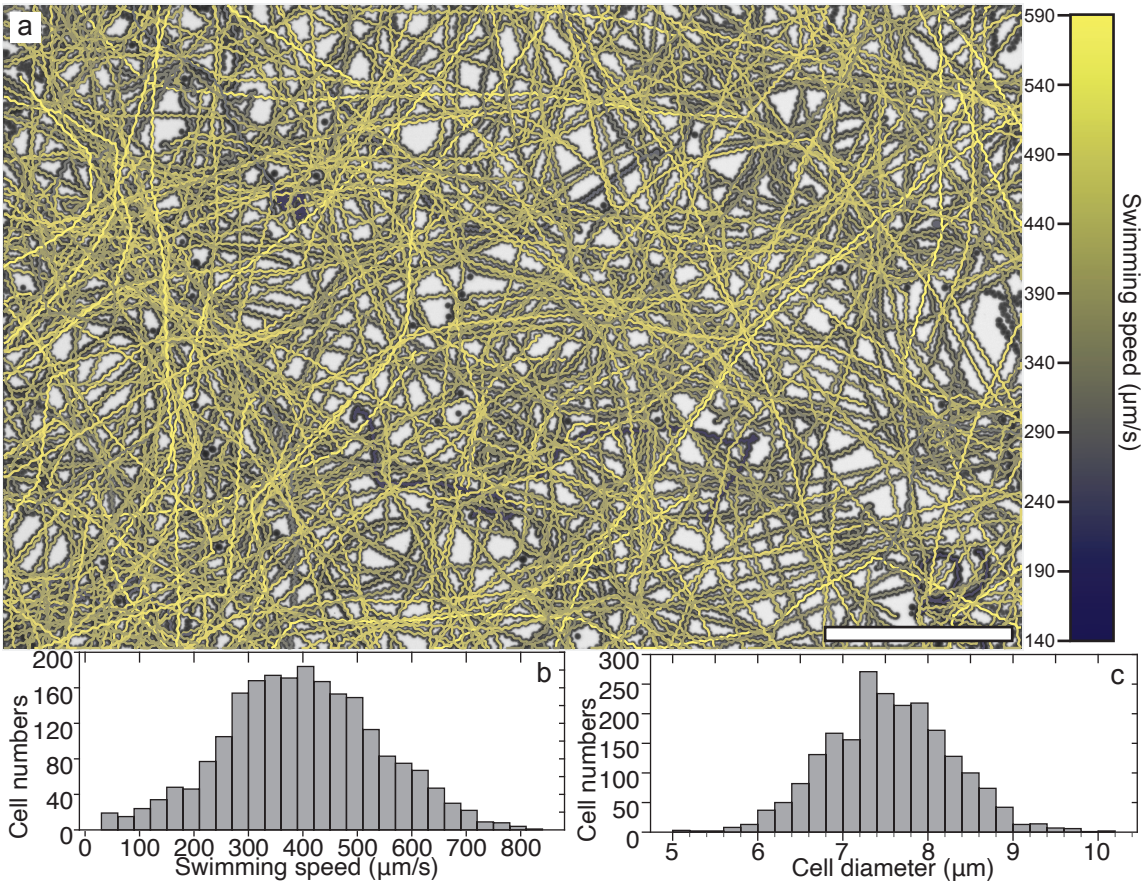


Figure 4 Measurement of swimming speeds for individual "*Ca.* T. celerrima" cells (**A**) One field of view showing all the motility tracks of "*Ca.* T. celerrima" cells captured during a 3 second, 364 frame acquisition. Tracks are colored by the average speed of the cell. Color bar in top right indicates speed, scale bar represents 200 μm. For representative video see **Movie S4**. (**B,C**) Histograms of cell swimming speed and diameter, respectively, accumulated from fourteen acquisitions from two separate wet mount preparations (2,147 individual cells).

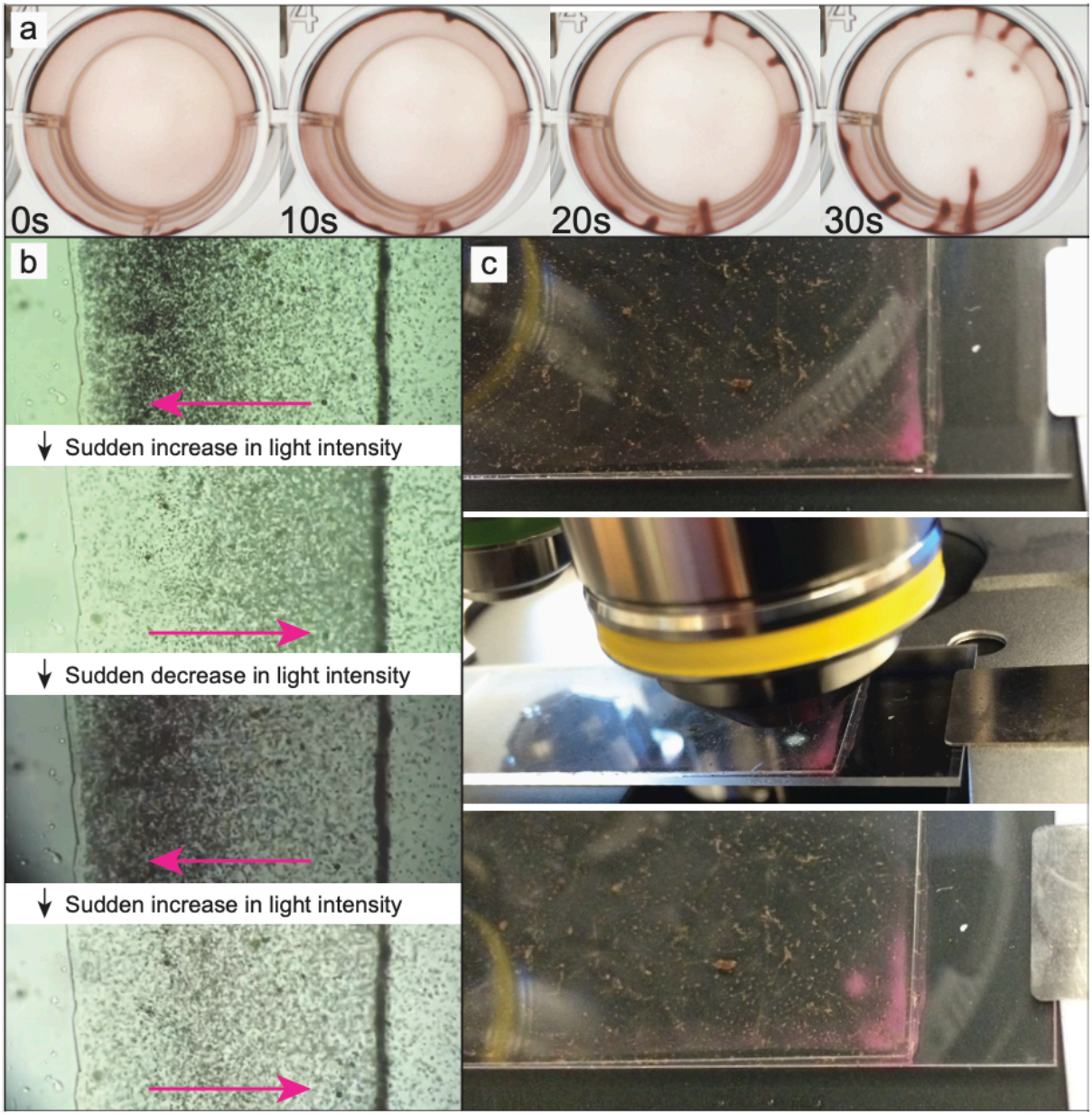


Figure 5 Light-based motility behaviors of "*Ca.* T celerrima". (**A**) Bioconvection was observed when dense preparations of "*Ca.* T celerrima" were mixed to homogeneity in a well of a 24-well plate containing sterile filtered source water. Over the 30 second time course represented here, cells concentrated themselves around the circumference of the meniscus and swarmed together towards the center of the well. Representative frames shown, for complete video see **Movie S5**. (**B**) Phobic motility response of "*Ca.* T celerrima" cells swarming at the edge of a wet mount to the rapid increase in illumination intensity. Representative frames shown, for complete video see **Movie S6**. (**C**) "*Ca.* T celerrima" cells accumulate in the constant illuminated region of a 10x objective over the course of approx. ten minutes. This response likely due to the classic "scotophobic" or "schreckbewegung" response.

	Radius (μm)	Speed (μm/s)	*Re*	*Pe*	Flow increase
Escherichia coli	0.25	16	4.00E-06	0.004	1.00
Chromatium okenii	2.875	27	7.76E-05	0.078	1.02
"*Ca* . T. celerrima" Free-swimming	3.745	399	1.49E-03	1.494	1.29
Thiovulum majus Free-swimming	5	600	3.00E-03	3.000	1.46
"*Ca* . T. celerrima" 1D Darcy's law	-	-	-	-	1.35
"*Ca* . T. celerrima" 3D Biofilm	-	-	-	-	1.37

Table 1 Substrate uptake enhancement by swimming. Comparison of "*Ca.* T celerrima" with *Escherichia coli*, a characteristic member of the Chromatiaceae family *Chromatium okenii*, and Thiovulum majus. The dimensionless Reynolds (*Re*) and Peclet (*Pe*) numbers are calculated from the cell size and swimming speed. *Re* values well below 1 demonstrate that viscous forces dominate over inertial forces in all cases. *Pe* numbers of ~1 for "*Ca.* T celerrima" under different assumptions and free swimming *Thiovulum majus* indicates qualitatively that these organisms exist at a boundary between advection and diffusion dominated mass transport regimes. Flow increase column represent quantitative estimates of the increase in diffusive substrate uptake due to advection. Cell size and swimming speed for non-"*Ca.* T celerrima" organisms are from ref 13. See Materials and Methods for details of calculations.

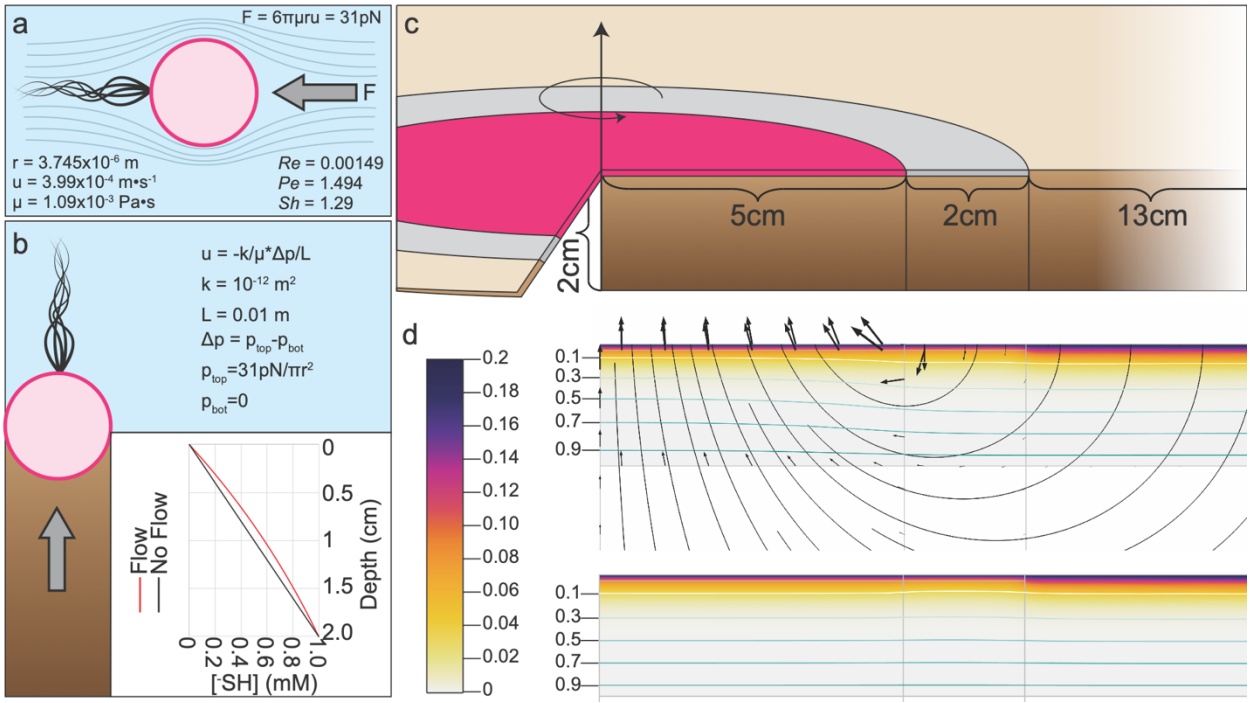


Figure 6 Theoretical evaluation of advection-based substrate uptake in "*Ca.* T celerrima". (**A**) Schematic of free-swimming "*Ca.* T celerrima" cell with parameters used in further modelling. Cell radius (r), average swimming speed (u) correspond to experimental data from **Figure 4**. Equation of force from Stoke's law is shown in the upper right. (**B**) Schematic of sediment-adhered "*Ca.* T celerrima" and parameters used for the calculation of sulfide flux via Darcy's law. (**C**) Schematic of radially symmetric 3D biofilm model including "*Ca.* T celerrima" and a surrounding biofilm of colorless sulfide oxidizing bacteria. Dimensions of the modeling domains are indicated. (**D**) representative modeling results illustrating flow lines (black lines) velocity vectors (black arrows) oxygen concentration (colored shading) and sulfide isopleths (colored lines). Upper and lower models are run with with and without "*Ca.* T celerrima"-induced pressure gradients, respectively.

Movie S1-3
Video microscopy of wet mounts of biofilms scraped from the side of polypropylene tubes containing "*Ca.* T celerrima" rapidly swimming and a variety of other associated organisms.
Movie S4
Representative example of 3 second video used for "*Ca.* T celerrima" tracking
Movie S5
Bioconvection of "*Ca.* T celerrima" in 24 well plate. In this video an enrichment of "*Ca.* T celerrima" is stirred to distribute the cells evenly throughout the well, and cells quickly accumulate at the edges of the well and begin bioconvection behavior.
Movie S6
Example of the effect of sudden increases in light intensity of the group motility of "*Ca.* T celerrima".
Movie S7
Example of "*Ca.* T celerrima" flagellar bundle rotation in a surface adhered cell
Movie S8
Example of two "*Ca.* T celerrima" beating their flagella while staying attached or tethered to the surface of the glass slide.

Supplementary file 1
Locus tags (IMG) of genes highlighted in the text as important for "*Ca.* T celerrima" motility or energy metabolism.
Supplementary file 2
Details of 1D modeling calculations and data for **Figure 6B**.
Supplementary file 3
Details of 3D modeling parameters, implementation and output.
Supplementary file 4
Data from 3D modeling used to evaluate the increase in substrate consumption due to flow.
Supplementary file 5
Full COMSOL output of 3D model.

Chapter 5

5. CONVERGENT EVOLUTION OF UNUSUAL COMPLEX I HOMOLOGS WITH INCREASED PROTON PUMPING CAPACITY: ENERGETIC AND ECOLOGICAL IMPLICATIONS

Grayson L Chadwick[1], James Hemp[1], Woodward W Fischer[1], Victoria J Orphan[1]

[1]Department of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91106

Abstract

Respiratory complex I is part of a large family of homologous enzymes that carry out the transfer of electrons between soluble cytoplasmic electron carriers and membrane-bound electron carriers. These complexes are vital bioenergetic enzymes that serve as the entry points into electron transport chains for a wide variety of microbial metabolisms, and electron transfer is coupled to proton translocation. The core complex of this enzyme is made up of 11 protein subunits, with three major proton pumping subunits. Here, we document a large number of modified complex I gene cassettes found in genome sequences from diverse cultured bacteria, shotgun metagenomics, and environmentally-derived archaeal fosmids all of which encode a fourth proton pumping subunit. The incorporation of this extra subunit into a functional protein complex is supported by large amino acid insertions in the amphipathic helix that runs the length of the protein complex. Phylogenetic analyses reveal that these modified complexes appear to have arisen independently multiple times in a remarkable case of convergent molecular evolution. From an energetic perspective, we hypothesize that this modification on the canonical complex I architecture allows for the translocation of a fifth proton per reaction cycle—the physiological utility of this modified complex is discussed.

Introduction

The generation and maintenance of an electrochemical proton gradient across the cytoplasmic membrane is centrally important for nearly all cells (Mitchell, 1966). The energy stored in this gradient is used for ATP generation as well as the active transport of a range of metabolites and other ions to and from the cytoplasm through the action of a wide array of symporters and antiporters. Membrane-bound electron transport in respiring organisms is carried out in order to produce and maintain this gradient. Complex I, so named due to its position in the mitochondrial electron transport chain, is one of the most broadly distributed electron transport complexes in biology. Complex I homologs are capable of interacting with widely utilized cytoplasmic electron carriers such as NADH, $F_{420}H_2$ (Bäumer et al., 2000), flavodoxins (Weerakoon and Olson, 2008), and likely ferredoxins (Battchikova et al., 2011; Welte and Deppenmeier, 2011), depositing their electrons onto many diverse membrane-soluble electron carriers including ubiquinone, menaquinone, and methanophenazine. Members of this family are often named based on the identity of the electron donors and acceptors such as NADH:ubiquinone oxidoreductase (Nuo), NADH:quinone oxidoreductase (Nqo), NADH dehydrogenase (Ndh), F_{420}:methanophenazine oxidoreductase (Fpo), and $F_{420}H_2$:quinone oxidoreductase (Fqo). This substrate versatility is due to the modular nature of the complex, where the electron input and output modules can be adapted or replaced to facilitate new interactions, while keeping the core function of redox powered proton pumping intact.

Understanding of complex I function has benefited greatly from recent high-resolution crystal structures from thermophilic bacteria and yeast mitochondria (Baradaran et al., 2013; Efremov et al., 2010; Hunte et al., 2010) as well as single-particle electron cryo-microscopy of mammalian mitochondria

(Vinothkumar et al., 2014). These studies reveal significant similarities shared between distantly related complexes (Sazanov, 2015). Complex I consists of three modules that are summarized briefly here. The Nuo naming scheme is used throughout which assigns the names NuoA-N to genes in the order they appear in the operon of the 14 subunit complex that oxidizes NADH and reduces ubiquinone. NuoEFG form the cytoplasmic N-module that interacts with NADH and facilitates its oxidation (**Figure 1A**). The NuoF protein binds an FMN cofactor, which serves as an intermediary between the 2-electron carrying NADH and 1-electron carrying iron sulfur clusters that transfer the electrons through the complex. The NuoG protein contains multiple iron sulfur clusters, and is homologous to members of the complex iron sulfur molybdopterin (CISM) superfamily of proteins, although it does not bind the bis(MGD)Mo cofactors generally required for catalytic activity in the CISM superfamily (Grimaldi et al., 2013).

In many *Epsilonbacteraeota* (Waite et al., 2017) such as *Campylobacter jejuni* and *Helicobacter pylori* the complex I operons contain NuoG but lack the NuoEF genes and have instead incorporated two unrelated protein subunits in their place (**Figure 1B**). These two subunits allow these complex I homologs to use a flavodoxin protein as the electron donor instead of NADH (Weerakoon and Olson, 2008). The N-module is absent in archaea where it is sometimes functionally replaced by FpoF—a subunit that binds an FAD cofactor and oxidizes a lower potential electron carrier, $F_{420}H_2$ (Bäumer et al., 2000; Friedrich and Scheide, 2000) (**Figure 1C**). FpoF is not often associated with the rest of the Fpo gene cluster. Many other bacterial and archaeal complex I operons also lack the N-module, with no replacement contained in the gene cluster, leaving the identity of the physiological electron donor in question (Moparthi and Hägerhäll, 2011) (**Figure 1D**). Hypotheses for the electron donor for these complexes include the recruitment of other proteins to functionally replace the diaphorase activity of the N-module,

as was hypothesized for the Hox hydrogenase (Schmitz and Bothe, 1996) and NdhV (Gao et al., 2016) in Cyanobacteria, FAD/NAD(P) binding oxidoreductases in ammonia oxidizing *Thaumarchaea* (Spang et al., 2012), or the complete absence of an N-module analog, and instead the direct interaction between the complex I homolog and soluble proteinaceous electron carriers such as ferredoxins (Friedrich et al., 1995).

The Q-module is comprised of subunits NuoBCDI that together bind additional iron-sulfur clusters and accept electrons from the N-module; they also partially facilitate the reduction of a quinone bound on the cytoplasmic face of the membrane. NuoD is homologous to the large catalytic subunits of NiFe hydrogenses but, in an interesting evolutionary parallel to NuoG, appears to have lost the residues required to bind the active site cofactors.

The remaining subunits NuoAHJKLMN are all integral membrane proteins referred to as the P-module. NuoL, M, and N pump one proton each during a reaction cycle, with a final fourth proton being pumped by the combined action of the remaining NuoAHJK subunits (Jones et al., 2017; Sazanov, 2015). NuoL, M and N are homologous to one another as well as to subunits from sodium-proton antiporters, formate-hydrogen lyases, carbon monoxide dehydrogenases, and membrane bound hydrogenases—relationships that reveal the shared evolutionary history of many major types of bioenergetic proton pumping complexes (Moparthi et al., 2014). NuoL contains a unique structural feature that sets it apart from NuoM and NuoN: a C-terminal extension that forms an amphipathic helix reaching from the far end of the complex back to the NuoN subunit (Sazanov, 2015)(illustrated in **Figure 1**). It is thought that this helix is important for transducing the energy released in the redox reactions of the Q-module into conformational changes in the P-module, leading to proton translocation. This helix has been

described as a molecular "piston" (Efremov et al., 2010), and site-directed mutagenesis studies creating insertions or deletions in this helix have deleterious effects on complex I assembly and function (Belevich et al., 2011; Torres-Bacete et al., 2011). The Fpo complex of methanogenic archaea likely pumps fewer protons based on biochemical measurements of the H^+/e^- stoichiometry, which are consistent with the less favorable energetics of the F_{420}/methanophenazine redox couple (Deppenmeier et al., 1990)(**Figure 1B**) and sequence analysis of the FpoL gene from a variety of methanogenic archaea (Welte and Deppenmeier, 2014).

Here we describe multiple occurrences and evolutionary histories of gene cassettes encoding complex I homologs that incorporate a second full-length copy of the NuoM subunit. These modified complex I homologs can be found in many phylogenetically diverse organisms, and appear to have arisen through convergent evolution involving the horizontal gene transfer of a second, distantly related copy of NuoM. Importantly, comparisons of the primary amino acid sequences of the peptides encoded by these gene cassettes supports the proposal that this extra NuoM is incorporated into a single functional complex by modifications that lengthen the NuoL amphipathic helix. The inclusion of the additional subunit suggests the upgrading of these complexes to a $5H^+/2e^-$ stoichiometry, which if correct would reflect the most protons pumped per reaction cycle by any bioenergetic complex.

Materials and methods

Complex I subunit gene selection

A rarified subset of microbial genomes (1,860) were selected from nearly 26,000 archaeal and bacterial genomes available on the IMG (Markowitz et al., 2012) server in April 2015. The rarified list was selected manually, keeping approximately one genome from each prokaryotic genus. Finalized genomes were

favored over partial or draft genomes when available. Initial complex I gene homologs were selected by retrieving all genes from this genome set automatically annotated as one of the Nuo genes by Kegg Ortholog number (see below). In our experience these proteins tend to be well annotated, although there are occasional contaminating proteins from other bioenergetics subunits, which were manually excluded from the analysis if they did not fall in what appeared to be a complex I operon. After identifying complex I clusters with additional pumping subunits, additional sequences from close relatives of those complexes were retrieved from databases including NCBI and IMG/M.

Protein sequence alignment

All proteins annotated as NuoL (including NuoL, NuoLM and NdhF [K00341, K15863, and K05577]) and NuoM (including NuoM, NuoLM, and NdhD [K00342, K15863, and K05575]) were separately aligned with Clustal Omega (Sievers et al., 2011), MUSCLE (Edgar, 2004) and MAFFT (Katoh et al., 2002). Many of the loops between transmembrane helices in NuoL and NuoM have lineage-specific indels that are not homologous features, and confound phylogenetic reconstruction, so we utilized GBlocks to identify well-conserved positions in our alignments that were suitable for building phylogenies (Talavera et al., 2007). The resulting alignments consisted of 274, 249 and 232 positions for NuoL from Clustal Omega, MAFFT, and MUSCLE, respectively. 201, 169 and 118 positions for NuoM, from Clustal Omega, MAFFT, and MUSCLE, respectively. 201, 154 and 145 positions for NuoN, from Clustal Omega, MAFFT, and MUSCLE, respectively. We further rarified the dataset using a 90% identity cut-off to reduce the dataset to a manageable size for phylogenetic reconstruction. After rarification the datasets contained between ~650 and ~1100 sequences depending on the subunit and alignment method used.

Phylogenetic reconstruction

For the large protein datasets, trees were constructed with RAxML for all three alignment methods. RAxML trees were built with RAxML-HPC v.8 on the CIPRES Science Gateway V 3.3 with 25 distinct rate categories, CAT model, with the LG substitution matrix. Only the Clustal aligned trees are shown in main text figures, those based on the other two alignment methods are shown in supplementary figures. Clade-specific trees constructed with the same RAxML settings, alignments were made with the three aforementioned alignment programs without GBlocks.

Structural homology modeling

NuoL genes from each clade and the NuoG gene from *Nitrospira defluvii* were submitted to the I-TASSER webserver (Roy et al., 2010; Yang et al., 2015; Zhang, 2008). For Clade I the NuoL sequence from *Chloroflexus* sp. Y-400-fl was submitted and the top model had a C-score of 0.4, TM-score 0.77+/-0.1 and RMSD of 7.3+/-4.2. For Clade 2 the NuoL sequence from *Desulfonatronovibrio magnus* was submitted and the top model had a C-score of 0.82, TM-score 0.82+/-0.08 and RMSD of 6.1+/-3.8. For Clade 3 the NuoL sequence from *Nitrospira defluvii* was submitted and the top model had a C-score of 0.62, TM-score 0.8+/-0.09 and RMSD of 6.7+/-4. The top model for *N. defluvii* NuoG had a C-score of 0.08, TM-score 0.72+/-0.11 and RMSD of 8.5+/-4.5.

Genome sequencing and assembly

The genomes of *Bellilinea caldifistulae* GOMI-1 (DSM 17877) and *Literilinea aerophila* PRI-4131 (DSM 25763) were sequenced as part of a project to expand the phylogenetic breadth of Chloroflexi genomes. Genomic DNA was ordered from Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ) and sequenced using the Illumina MiSeq platform. SPAdes 3.1.1 (Nurk et al., 2013) was used

to assemble the genomes. For *B. caldifistulae*, sequence coverage, GC composition and phylogenetic affiliation of conserved single copy genes were used to exclude contaminating contigs and genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline. The draft genome is 3.72 Mb in size, comprised of 31 contigs, 3399 genes, 2990 CDSs, and is estimated to be ~95% complete based on conserved single copy genes (106/111). This whole-genome shotgun project has been deposited in DDBJ/EMBL/GenBank under the accession number LGHJ00000000. *L. aerophila* sequencing data was not high enough quality to produce a draft genome, however, SPAdes assembly did produce a contig containing the complex I gene cassette. The DNA sequence for this contig is deposited in NCBI under accession number JWSX01002273.1.

Results and discussion

Diverse complex I gene cassettes with extra proton pumping subunits

Through detailed investigation of the diversity of complex I homologs we discovered a unique type of complex I gene cassette encoding a second full-length copy of the large proton pumping subunit NuoM. These gene clusters (herein referred to as 2M complexes) are observed in genomic data from five different bacterial phyla, environmentally-derived *Thaumarchaeal* fosmids, and a number of genome bins from assembled metagenomes. Examples of 2M complex gene clusters are shown in **Figure 2A**, and were initially classified into three clades according to their gene synteny and primary sequence similarity. Phylogenetic trees built from NuoL homologs from a broad sampling of publically available data reveal that the three clades of 2M complexes form phylogenetically coherent groups comprised of organisms from: Clade 1) *Chloroflexi*, *Verrucomicrobia* and *Acidobacteria*; Clade 2) uncultured *Thaumarchaea* and *Deltaproteobacteria* of the *Desulfovibrionales* order; and Clade 3) the genus *Nitrospira* (**Figure 2A**). In each case the most closely related NuoL proteins compared to those found in the 2M clades occurred in

canonical complex I operons without the second NuoM subunit (as in **Figure 2A**). This suggests three separate evolutionary events in which a second NuoM was added to a canonical complex I operon. The accuracy of these phylogenetic relationships were evaluated using two different primary sequence alignment tools—both yielded phylogenies confirming the polyphyletic nature of the 2M clades (Methods, **Figures S1-3**)

To determine the origins of these second NuoM subunits contained in the 2M operons we constructed phylogenetic trees from a broad sampling of NuoM subunits in a similar manner as NuoL. As was observed with the NuoL phylogeny, all three clades were distinct from one another in the NuoM tree (**Figure 3A**). However, the two copies of NuoM from each clade are not closely related to one another. This illustrates that the second NuoM was not derived from direct duplication of NuoM from within the operon, but instead was incorporated from a separate complex I homolog through horizontal gene transfer. Interestingly the NuoM2 gene from the Clade 3 *Desulfovibrionales* is located within a completely different region of the tree than the NuoM2 from the *Thaumarchaeal* fosmids of this clade. It appears that although the NuoL subunits in Clade 3 are all closely related, the additional NuoM subunits were acquired independently in these two separate lineages. Phylogenies of the last large proton pumping subunit NuoN revealed the same polyphyletic pattern of the three 2M clades, including the separate clusters between the *Desulfovibrionales* and *Thaumarchaea* members of Clade 3. Additional NuoM and NuoN trees were constructed using the same approach as with the NuoL trees, all yielded similar relationships (**Figures S4-9**).

Evolution of 2M gene cassettes

To increase our understanding of the evolution and diversity of 2M complexes, we retrieved many additional members of the 2M clades to build clade-specific phylogenies. We also retrieved their nearest relatives from the NuoL tree in **Figure 2A** to use as outgroups for our analysis. These additional 2M gene cassettes came from publically available data as well as our own sequencing efforts targeting diverse groups of cultured *Chloroflexi* (see Methods). The operon structures of this expanded collection of 2M complexes are shown in **Figure 4A-C** and are ordinated by the phylogenetic relationships of their NuoL proteins. NuoL was chosen for reconstructing their phylogenetic history because it is the largest protein in the operon (~600 amino acids) and there is reason to believe that concatenating all of the genes together to make a composite phylogeny would be inappropriate due to the possibility that the individual genes in these operons have distinct evolutionary histories. In all three clades there is a high degree of congruence between the NuoL phylogeny and gene synteny in the operons suggesting the tree topology accurately captures key aspects of the evolution of these gene clusters.

The 2M clades differ from one another in their operon structure and gene content. All members of Clades 1 and 3 have inserted the additional NuoM2 subunit between the original NuoL and NuoM1 subunits (**Figure 4A,C**), whereas the representatives from Clade 2 all have the gene order NuoMLMN, with the additional NuoM2 gene being inserted between the L and N (**Figure 4B**). The gene order with a NuoM preceding NuoL is characteristic of *Thaumarchaeal* and *Crenarchaeal* complex I operons, which agrees with the phylogenetic placement of this group with the NuoL from canonical complex I operons in these archaea, as well as some members of Clade 2 being found in *Thaumarchaeal* fosmids (Deschamps et al., 2014).

In Clade 1 three large subgroups are apparent based on the phylogeny of NuoL, the phylogeny of the host organism, and the operon structure (**Figure 4A**). Clade 1a contains examples found in *Verrucomicrobia* and *Acidobacteria*, Clade 1b contains members of the *Chloroflexia* class of the *Chloroflexi* and Clade 1c is found in *Thermoflexia* and *Anaerolineae* classes of the *Chloroflexi*. All members of Clade 1 lack the N-module genes and exhibit a unique gene order where all subunits of the Q-module (NuoBCDI) occur one after the other. This pattern never occurs in normal complex I operons, and might be a result of the gene rearrangement(s) that led to the incorporation of an additional NuoM gene. Notably, the ORFs marked in red in the *Didymococcus colitermitum* TAV2 operon are annotated as an IS200-like transposase and a member of the PD-(D/E)XK nuclease superfamily, which could be remnants of the mobile genetic elements responsible for the genome rearrangements that formed these 2M operons (**Figure 4A**). In Clade 1c the NuoBCDI genes are found together in a separate part of the genome, with NuoC and NuoD fused (as depicted in **Figure 2A**). This NuoCD fusion is not unique to 2M complexes and also exists in normal complex I operons (Friedrich, 1998), but the separation of the Q-module genes from the rest of the complex is peculiar.

Clade 3 operons from the *Nitrospira* genus have two important differences between them and the other 2M complexes. First, is the complete absence of a NuoH gene. Members of the *Nitrospira* also have a canonical Nuo-type complex I however, so at least one copy of NuoH is present in these organisms; perhaps this NuoH is used by both complexes. No lone NuoH genes could be found elsewhere in any of the genomes. The second major difference is the presence of a large NuoG homolog in the gene cluster where the N-module proteins are normally located. As mentioned above, NuoG is part of the CISM superfamily, and in traditional complex I lacks the bis(MGD)Mo cofactor. However, the NuoG homolog in these gene cassettes is much more similar to members of the CISM superfamily than normal

NuoG proteins, particularly the large catalytic subunits of formate dehydrogenases, so it is possible that in the Clade 3 2M complexes the NuoG has a novel catalytic role. This similarity to formate dehydrogenases extends to the incorporation of an additional iron sulfur cluster, which connects the iron sulfur clusters used by canonical NuoG to the ancestral CISM active site as has been previously reported for the NuoG proteins of *H. pylori* and *C. jejuni* (Efremov and Sazanov, 2012). However, one of the key ligand binding residues, SeCys/Cys140 (PDB:1FDO numbering) used for ligating the molybdenum atom in the bis(MGD)Mo cofactor (Cerqueira et al., 2015), is replaced by a glycine in the Clade 3 NuoG homologs (**Figure S10**). The absence of this cysteine argues against the Clade 3 NuoG homologs being catalytically active in the known CISM fashion. The *H. pylori* and *C. jejuni* complex I gene cassettes lack NuoEF as well, but they have been replaced by two additional proteins that have resulted in the utilization of flavodoxin instead of NADH (Weerakoon and Olson, 2008). Neither of these two proteins have homologs in the *Nitrospira* genomes.

Extension of the NuoL amphipathic helix

The acquisition of a second copy of NuoM and its incorporation into a consistent well-ordered operon suggests that these 2M gene cassettes may produce larger versions of complex I with the ability to pump an extra proton reaction cycle. To further investigate this idea, we closely examined the sequences of the NuoL genes from these 2M complexes in light of available crystal structure data. If these 2M complexes contain an additional NuoM subunit one would expect the entire P-module to be extended in length by the width of this subunit, and a canonical NuoL subunit positioned at the distal end of a 2M complex would not be able to reach its amphipathic helix past the second copy of NuoM all the way to the NuoN (see **Figure 1**).

Remarkably, the alignments of the NuoL proteins in each of the three clades described above have an insertion of 26-28 amino acids in the middle of the amphipathic helix compared to their closest relatives lacking the second NuoM, as well as the NuoL sequence from *T. thermophilus* (**Figure 5A-C**). Although estimates of the exact position of these insertions vary slightly based on the sequence alignment algorithm used, they all place an insertion of this size within the amphipathic helix (see Methods, **Figures S11-13**). Using the *T. thermophilus* crystal structure, we measured the distance spanned by 26 amino acids along the amphipathic helix to be *ca.* 37 Å—a distance that nearly exactly matches the width of the NuoM subunit (**Figure 6A**). Structural homology models built using I-TASSER (Roy et al., 2010; Yang et al., 2015; Zhang, 2008) fit the NuoL subunits from all three 2M clades very well onto existing complex I crystal structures and have placed the extra 26-28 amino acids as insertions in the middle of the amphipathic helix in agreement with the multiple sequence alignments (**Figure 6B-D**).

Some of the closest relatives of the 2M complex NuoL genes contain the amphipathic helix insertion, but lack the additional NuoM. In Clade 1, *Caldilinea aerophilum*, *Ardenticatena maritima*, *Litorilinea aerophila*, and five members of the *Actinobacteria* outgroup complex all have the insertion (**Figure 5A**). No close relatives of Clade 2 exhibited this insertion, but the closest relative of the Clade 3 complexes, *Ca.* Methylomirabilis oxyfera, had an insertion without a second NuoM as well (**Figure 5C**). From this we interpret the insertion of the 26-28 amino acids in the NuoL amphipathic helix as a potentiating mutation that enabled the addition of a second, distantly related NuoM subunit. It is worth noting that many of the *Actinobacteria* (*Ca.* Microthrix parvicella, *Nitriliruptor alkaliphilus*, AAA007-J07 and AC-312-N20), *C. aerophila*, and *Ca.* M. oxyfera all have a second complex I gene cassette immediately up or downstream of the one depicted in **Figure 5**, raising the possibility that these organisms might use an additional NuoM from the second complex *in lieu* of forming their own 2M operon.

Bioenergetic considerations and physiological conditions leading to increased proton pumping

The physiological and ecological utility of these novel 2M complexes remains to be discovered, however multiple instances of their convergent evolution, combined with their diversity and cross-domain horizontal gene transfer suggests that they have meaningful adaptive value. We speculated that by containing an extra proton-pumping subunit, 2M complexes achieve a higher stoichiometry of protons translocated per $2e^-$ reaction cycle. On its face, such a trait might seem like an advantage for all metabolisms that employ complex I, but to inform this idea it is useful to examine the energetics of the reactions facilitated by these enzymes. From a thermodynamic point of view, complex I and its homologs convert energy released from a redox reaction into transmembrane proton motive force. Below we examine how the redox reaction and electrochemical proton gradient can influence the number of protons pumped per reaction.

The Gibbs free energy associated with the redox chemistry carried out at complex I is described by equation (1), which is simply the number of electrons transferred, multiplied by Faraday's constant and the potential of the redox couple (E). The potential of the redox couple is described by the Nernst equation (2), which is dependent on the standard state of the redox couple (E°), less the natural log of the ratio between the concentration of oxidized and reduced versions of each species.

(1) $\Delta G = -n_e F E$

(2) $E = E° - \frac{RT}{F} \ln K,\ K = \frac{[A][QH_2]}{[AH_2][Q]}$

Complex I homologs convert the free energy of these redox reactions into free energy stored in the proton transmembrane electrochemical gradient. This form of Gibbs free energy depends on the

number of protons translocated (n_H), multiplied by the difference in electrochemical proton potential across the membrane (equation (3)). The electrochemical potential is in turn dependent on the pH and charge gradients across the membrane (equation (4)).

(3) $\Delta G = -n_H \Delta\tilde{\mu}_H$

(4) $\Delta\tilde{\mu}_H = F\Delta\psi - 2.3RT\Delta pH$

Combining the two Gibbs free energy terms yields equation (5), which gives the free energy change associated with the combined reactions of electron transfer and proton translocation.

(5) $\Delta G = -n_H(F\Delta\psi - 2.3RT\Delta pH) - n_e F(E° - \frac{RT}{F}\ln K)$

Of particular interest for the discussion of these 2M complexes is the n_H term. Solving equation (5) for n_H yields equation (6) which illustrates how each of the variables affects n_H.

(6) $n_H = \frac{\Delta G + n_e F(E° - \frac{RT}{F}\ln K)}{2.3RT\Delta pH - F\Delta\psi}$

Many different aspects of the reaction catalyzed by complex I homologs can affect the number of protons pumped, but only a few seem likely to require an increased proton pumping capacity during electron transport in the organisms we have identified that contain 2M complexes. For example, a change in the identity or redox state of electron carriers may result in an increase in E°: *i.e.* a larger redox potential difference between the cytoplasmic electron carrier and the membrane bound electron carrier. Due to the positive dependence of E° on n_H in equation (6), if all other factors remain the same, this change would increase the number of protons that can be pumped. Similarly, decreasing ΔpH will increase the value of n_H, since there is an inverse relationship between these two values. Restating this in physiological terms: as the extracellular pH increases, more protons can be pumped per reaction since there is a weaker proton motive force resisting the action.

The effect ΔpH can have on enzymes relying on proton motive force has already received much attention in the context of alkaliphilic bacteria such as *Bacillus pseudofirmus* OF4. These organisms have adapted to high pH environments in part by increasing the number of c-ring subunits in their ATP synthases (Hicks et al., 2010; Meier et al., 2007; Preiss et al., 2014, 2013). This adaptation increases the number of protons translocated per ADP phosphorylated, and helps compensate for the fact that each proton translocation carries with it less energy. We hypothesize that in some cases the addition of the second NuoM subunit in 2M complexes could be a complex I equivalent of increasing the number of c-ring subunits in the *Bacillus* ATPase (**Figure 7A**). With respect to this hypothesis, it is worth noting that of the five *Desulfovibrionales* encoding 2M complexes, four are alkaliphiles with pH optima between 9.5 and 10 (Sorokin et al., 2011, 2008).

In the case of alkaliphiles described above, the environment imposes a decrease in the electrochemical proton gradient. Another possible reason for an organism to operate at an electrochemical proton gradient below "normal" levels may be the adaptation to slow growth. Protons can leak directly through the cytoplasmic membrane in a way that is decoupled from energy conservation. For organisms that metabolize catabolic substrates very slowly, this leakage may present a large problem to their viability (Niftrik et al., 2004; Valentine, 2007). The amount of proton leakage through a membrane is related to the electrochemical proton gradient across it (Rupprecht et al., 2010). In fact, the leakage current through the membrane appears to be non-linear with respect to transmembrane voltage, meaning a relatively small drop in transmembrane electrochemical proton gradient could result in a quite significant decrease in leakage current (Rupprecht et al., 2010). If the catabolic activity of a slow growing organism respiring at low rates cannot keep up with the leakage current experienced by fast-growing organisms at

high respiration rates, then one possible solution would be to employ biochemical adaptations that decrease their normal operating transmembrane potential. Living at this lower potential would require the pumping of more protons to conserve the same amount of energy contained in a redox reaction. The addition of an additional M subunit in complex I may help accomplish this.

It is noteworthy that many of the organisms highlighted here, particularly *Nitrospira* and ammonia oxidizing *Thaumarchaea*, are notorious for slow growth, low biomass yields, and substrate limited growth conditions (Daims et al., 2015; Ehrich et al., 1995; Könneke et al., 2005; Stieglmeier et al., 2014). Additionally, three of the *Verrucomicrobia* (TAV1, TAV2 and TAV5) and *Acidobacteria sp.* KBS 89 were all isolated in a single study which described a novel cultivation strategy specifically designed for slow growing organisms (Stevenson et al., 2004). Enrichments were incubated for more than 30 days and the colonies that were picked were invisible to the naked eye. Interestingly, the *Verrucomicrobia* were isolated from the termite gut, and the *Acidobacteria* from soil, but both were isolated under slow growth conditions and both contained a version of the 2M complex.

Finally, complex I can run in the reverse direction, producing low potential electron carriers (like NADH) from high potential donors (like ubiquinonol) at the expense of proton motive force. This reverse electron transport occurs in many chemolithoautotrophs and anoxygenic photoautotrophs, which utilize electron donors that are at a higher redox potential than the cytoplasmic electron carrier required for carbon fixation (Aleem, 1966). The *Nitrospira* that comprise Clade 2 (Daims et al., 2015; Koch et al., 2015) and the *Thaumarchaea* that comprise Clade 3 (Deschamps et al., 2014) both face this challenge with their autotrophic pathways. The ammonia-oxidizing *Thaumarchaea* utilize the highly efficient hydroxypropionate/hydroxybutyrate cycle for carbon fixation that requires electrons from

NADPH (Könneke et al., 2014), while *Nitrospira* utilize the reverse tricarboxylic acid cycle for carbon fixation, which requires the use of even lower potential ferredoxins (Lücker et al., 2010). In both cases reverse electron transport must occur, since the midpoint potential of NADPH (-320 mV) and ferredoxins (*ca.* -500 mV) are well below the potential of the NH_4^+/NO_2^- redox couple used for their energy metabolism (+340 mV)(Simon and Klotz, 2013). It may be that in these organisms, the 2M complexes are used specifically for reverse electron transport onto NADPH or ferredoxin (**Figure 7B**). In *Nitrospina gracilis* it has been proposed that a complex I operon lacking a traditional N-module could be responsible for reverse electron transport to the level of ferredoxin (Lücker et al., 2013). Close relatives of that version of complex I do not appear in the Nitrospira genomes recovered to date, but the 2M complexes described here could in principal provide a similar energetic advantage under conditions of reverse electron transport. Reverse electron transport may be an important function of the 2M complexes in photosynthetic *Chloroflexi* as well. These organisms use reaction center II for cyclic electron flow, which, unlike reaction center I, does not produce a reduced ferredoxin (Fischer et al., 2016). Therefore, to withdraw electrons from the photosynthetic electron transport chain and transfer them to lower potential electron acceptors like NAD(P)H for carbon assimilation and anabolism, there must be an input of energy for reverse electron transport from quinols. It is possible that in these organisms reverse electron transport can only be achieved using the increased driving force of the 2M complexes.

Conclusions

Complex I and its homologous bioenergetic complexes have been modified and repurposed many times during the evolution of life on Earth. This evolution has involved the addition and removal of subunits specific for certain substrates onto a membrane-integral chassis containing between 1 and 3 large proton

pumping subunits (Moparthi et al., 2014). The 2M complex I homologs described here represent a fascinating modification on the canonical complex I enzyme that incorporate a fourth large pumping subunit. The phylogenetic analyses presented here demonstrate that these 2M complexes are polyphyletic, with respect to both the phylogeny of the host organism and the evolutionary history of the individual major pumping subunits. Combined with the unique gene synteny found in each of the three clades this strongly supports the conclusion that the construction of these operons resulted from multiple cases of convergent molecular evolution.

The highly conserved nature of the 2M complexes in each of the three respective clades implies that these are not pseudogenes or evolutionary dead-ends. The construction of gene clusters with the additional subunit, as well as the precise size and location of the insertions in the amphipathic helix of NuoL implies that this increase in pumping potential has been selected for under specific adaptive pressures. Future work is required to understand the conditions under which the complexes are expressed, and verify their subunit composition, substrate specificity, and proton pumping stoichiometry. However, the physiological diversity of the organisms containing these complexes suggests that they have evolved in response to different selective pressures and for different functions. The evolution of these complexes appears to have involved changes to the genetic content of an organism through horizontal gene transfer, genome restructuring through operon rearrangement, and protein primary, secondary, tertiary and quaternary structure modification in the extension of the NuoL amphipathic helix. Altogether, these observations highlight the different biological scales over which changes occur in order to bring about functional evolutionary transitions.

Figures

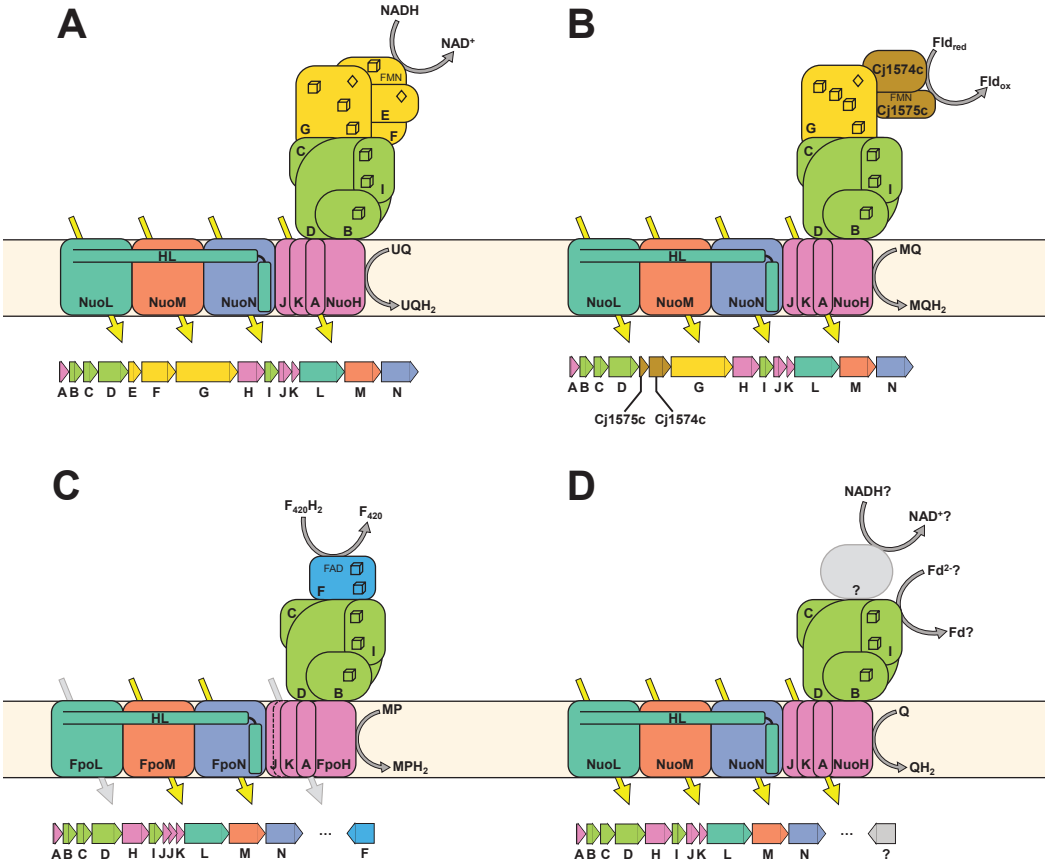


Figure 1 Common complex I homolog quaternary structure and gene order. Yellow arrows indicate proton translocations from the cytoplasm to the periplasm through the membrane-bound P-module. The C-terminal amphipathic helix of NuoL is labeled HL. 4Fe4S sulfur clusters are depicted as cubes, 2Fe2S clusters as diamonds. (**A**) Canonical 14 subunit NADH:ubiquinone oxidoreductase (Nuo) as found in *T. thermophilus*. Genes are generally arranged in an operon in alphabetical order as depicted. (**B**) A complex I cassette found in *Epsilonbacteraeota* and best characterized in *C. jejuni* and *H. pylori*. The NuoEF genes are replaced by two smaller, non-homologous, proteins, which facilitate the interaction with flavodoxin (Fld) instead of NADH. The NuoG proteins in these operons contain an extra iron sulfur cluster. (**C**) A methanogen-type $F_{420}H_2$:methanophenazine oxidoreductase (Fpo) showing the replacement of the N-module (NuoEFG) with the FpoF gene (in blue). This gene is not commonly associated with the rest of the operon, but its inclusion in the functional protein complex has been demonstrated by purification of the native Fpo complex. Note: two genes are annotated as NuoJ in these operons, but this represents a fission of the canonical NuoJ gene, and not a gene duplication event. (**D**) A common complex I gene cassette found in many organisms containing only the core 11 subunits of the complex. Few of these complexes have been studied biochemically and it is currently unknown whether the electron donors for these complexes are soluble proteinaceous electron carriers such as ferredoxins (Fd), or whether additional proteins encoded elsewhere in the genome are recruited to the complex to interact with small molecule electron carriers like NADH.

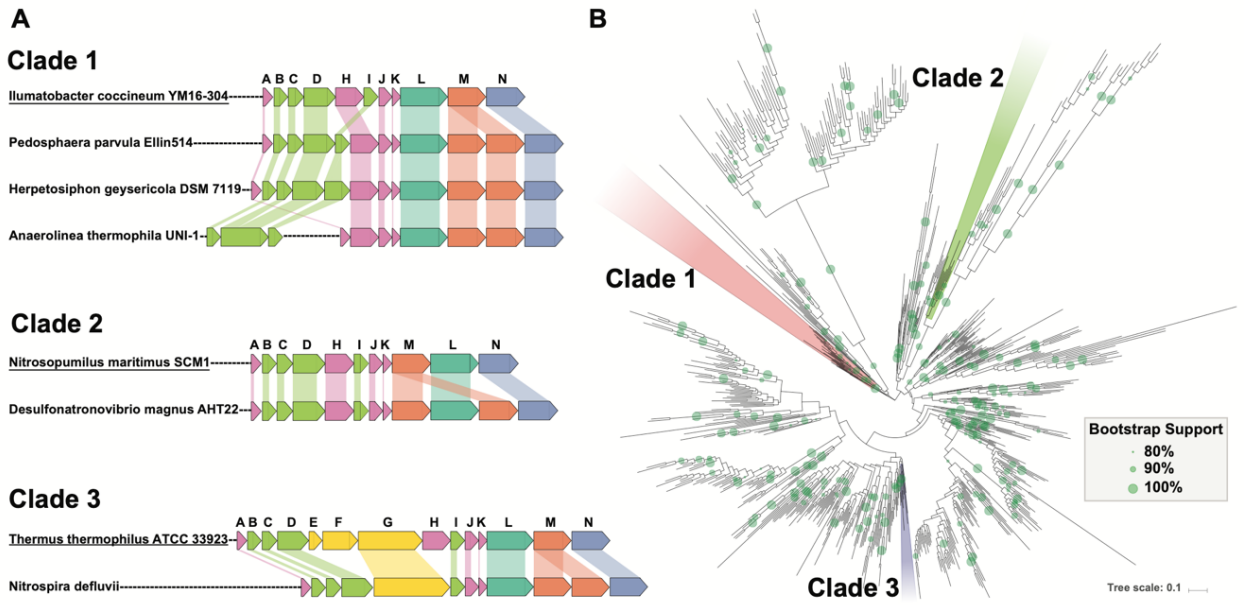


Figure 2 2M complex I homologs were found in three discrete clades. (**A**) Examples of the operon structure of 2M complexes from each of the three different clades showing the addition of a second NuoM subunit (in orange) as well as general operon features. Closely related canonical complex I operons are shown for comparison (underlined). In Clade 1 the second NuoM occurs right after the first, and the operons contain a unique gene rearrangement that swaps the NuoH and NuoI subunits; in some members of this Clade such as *Anaerolinea thermophila* UNI-1 the NuoBCDI genes occur separately from the rest of the operon. In Clade 2 the two NuoM genes are found on either side of NuoL; the NuoM preceding the NuoL is characteristic of the complex I operons found in Thaumarchaea. None of the operons have full N-modules, and only Clade 3 has a NuoG homolog in its operon. (**B**) A maximum likelihood phylogenetic tree of NuoL subunits from a rarified subset of publically available genome data. The NuoL genes from Clades 1, 2 and 3 fall far from one another in well-supported groups surrounded by NuoL genes in normal complex I operons—a pattern that implies that multiple independent events of convergent evolution gave rise to the 2M complex Is. Bootstrap support is shown for values ≥80%. High resolution NuoL trees, including those built from MUSCLE and MAFFT alignments are shown in Figures S1-3.

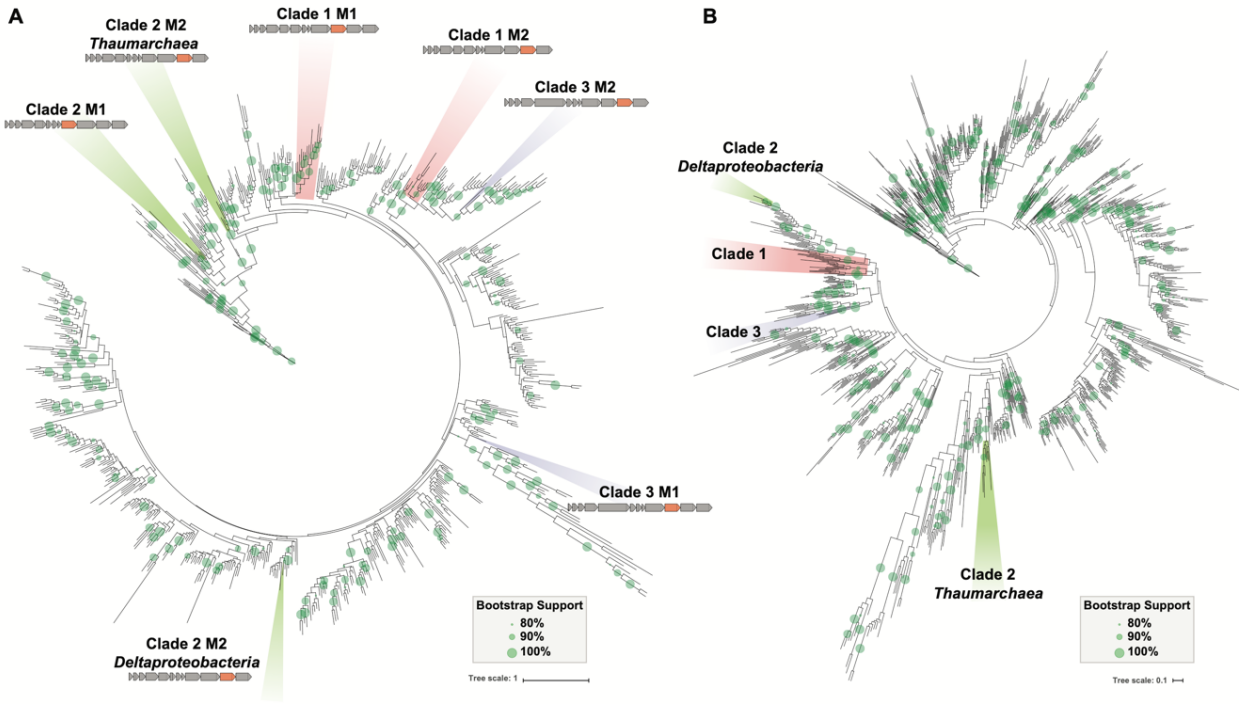


Figure 3 Phylogenetic trees of NuoM and N subunits from a rarified subset of publically available genome data. Bootstrap support is marked on the branches with values ≥80% support are shown. (**A**) As observed in the NuoL tree, the NuoM genes from Clades 1, 2 and 3 remain separate and distinct. Within Clade 2 the NuoM2 genes are appear to be polyphyletic, with separate origins for the *Thaumarchaeal* NuoM2 and the *Deltaproteobacterial* NuoM2. Operon diagrams illustrate which NuoM homolog is found in each position of the tree. (**B)** The NuoN genes also show the polyphyletic nature of the three 2M clades, as well and separation between *Deltaproteobacteria* and *Thaumarchaeal* members of Clade 3. High resolution NuoM and N trees, including those built from MUSCLE and MAFFT alignments are shown in Figures S1-3.

Figure 4 2M Clade-specific NuoL phylogenies with operon structures. (**A**) Clade 1 phylogeny reveals three major groups; 1a: *Verrucomicrobia* and *Acidobacteria*, 1b: *Chloroflexia*, 1c: *Thermoflexia* and *Anaerolinea*. All members and outgroups for Clade 1 lack NuoEFG genes and exhibit the unique reordering of the NuoH/I genes as highlighted in Figure 2a. Possible mobile genetic elements in the *Didymococcus colitermitum* operon are highlighted in red. (**B**) Clade 2 is comprised of members of the *Deltaproteobacteria* and marine *Thaumarchaea* fosmids; it shares the NuoML gene order found in the other known *Thaumarchaea* used as an outgroup. (**C**) Clade 3 is comprised of members of the *Nitrospira* genus and contains a large NuoG gene as the only subunit of the N-module. The designations NuoM1 and NuoM2 were assigned based on which copy was more similar to the NuoM found in the most closely related normal complex I operons. Note: some operons are truncated due to their location on the ends of genomic contigs, and do not reflect segmentation or truncation in intact genomes.

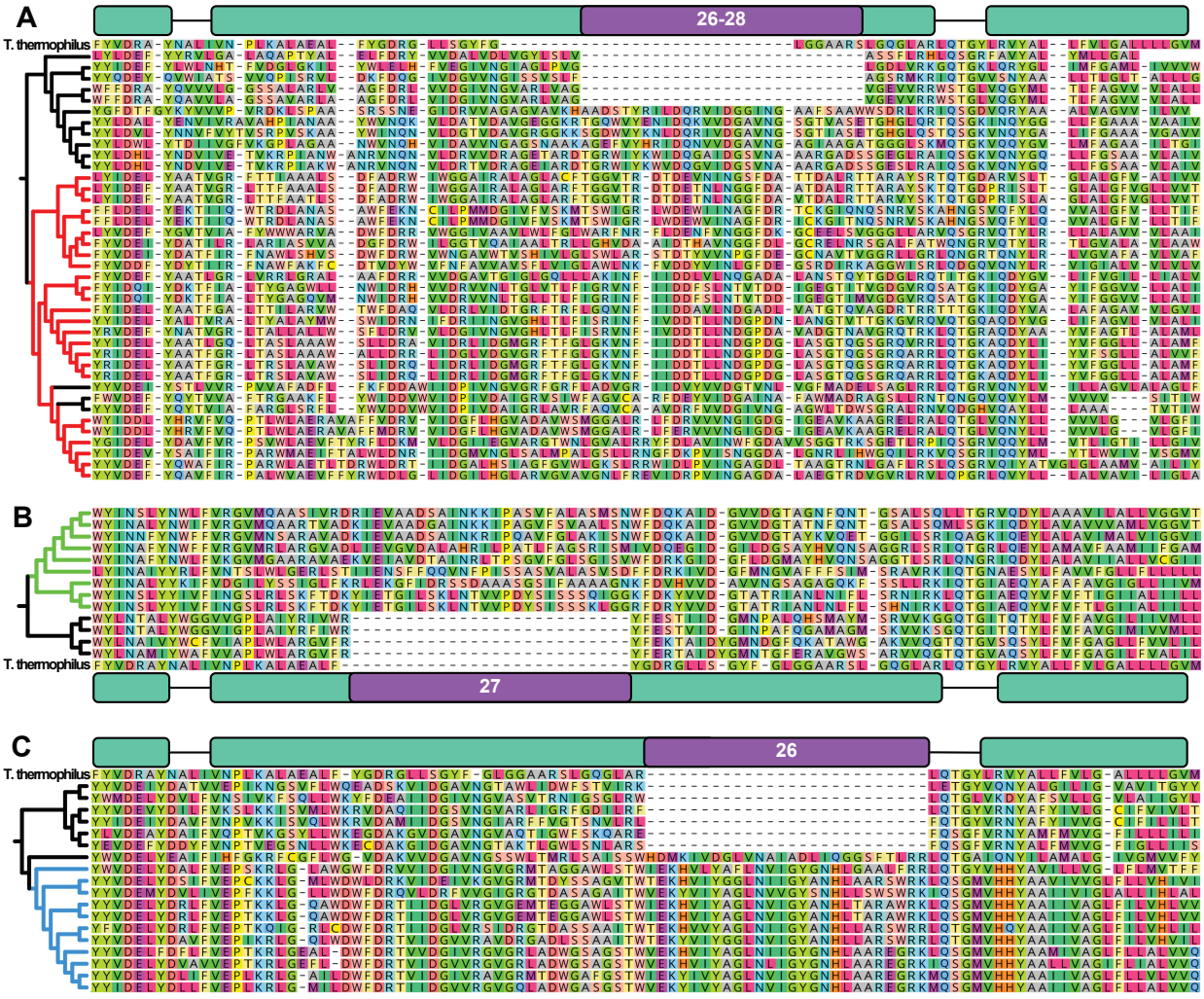


Figure 5 Sequence alignments highlighting the amphipathic helix insertions of the NuoL subunits from 2M complexes and their close relatives. Cladograms represent the same tree topology as those in **Figure 4**, with colored branches marking the 2M groups. The *T. thermophilus* sequence is included in the alignment as a sequence length metric of where the amphipathic helix begins and ends. (**A**) Clade 1 contains 26-28 amino acids insertions compared to many of the close relatives and *T. thermophilus*. (**B**) Clade 2 contains 27 amino acid insertions compared to their closest relatives (28 compared to *T. thermophilus*). (**C**) Clade 3 contains 26 amino acid insertions. Notably, in (**A**) and (**C**) the closest relatives of the 2M complexes that were used as outgroups in the phylogenetic analysis also have insertions in the amphipathic helix, despite not having a second distantly related NuoM in their complex I operons. MUSCLE alignments are shown here, Clustal Omega and MAFFT alignments are shown in Figures S11-S13.

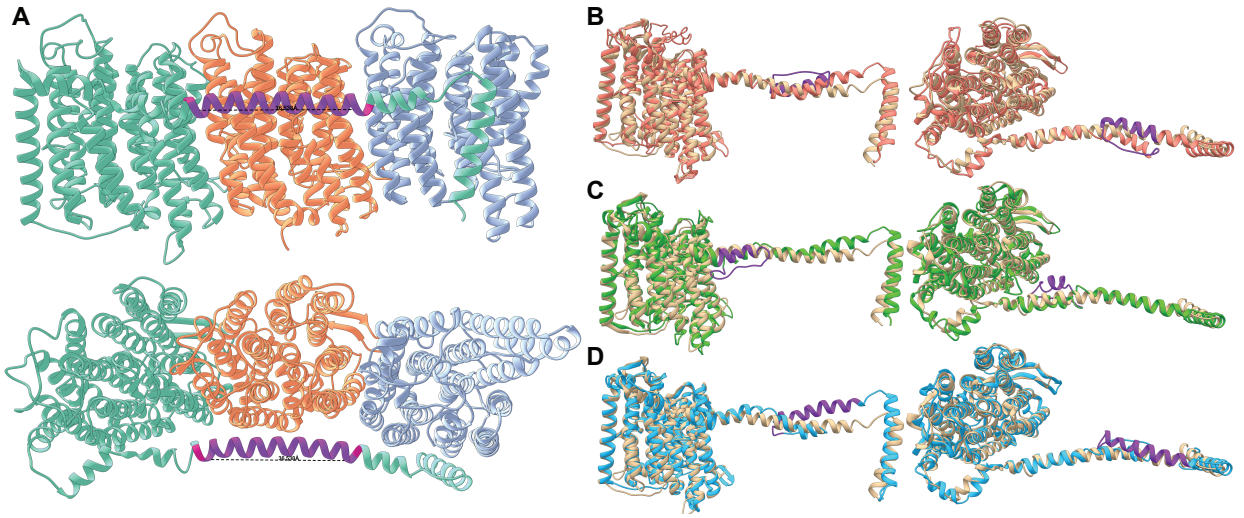


Figure 6 Nuo crystal structure and structural homology models highlighting the characteristic amino acid insertion lengthening the amphipathic helical arm that spans the 2M complexes. (**A**) NuoL, M and N subunits from the 4HEA crystal structure with a stretch of 26 amino acids marked in purple on the amphipathic helix (28 denoted by the additional red residues on either end of the purple region). This length of helix is approximately the same width as the NuoM subunit shown in orange. I-TASSER structural homology models of NuoL from Clades 1, 2, and 3 modeled onto the NuoL subunits from 4HEA are shown in (**B**), (**C**), and (**D**), respectively. In all cases the insertions noted in Figure 5 have been colored purple, and were modeled by I-TASSER as helices that doubled back on the original helix.

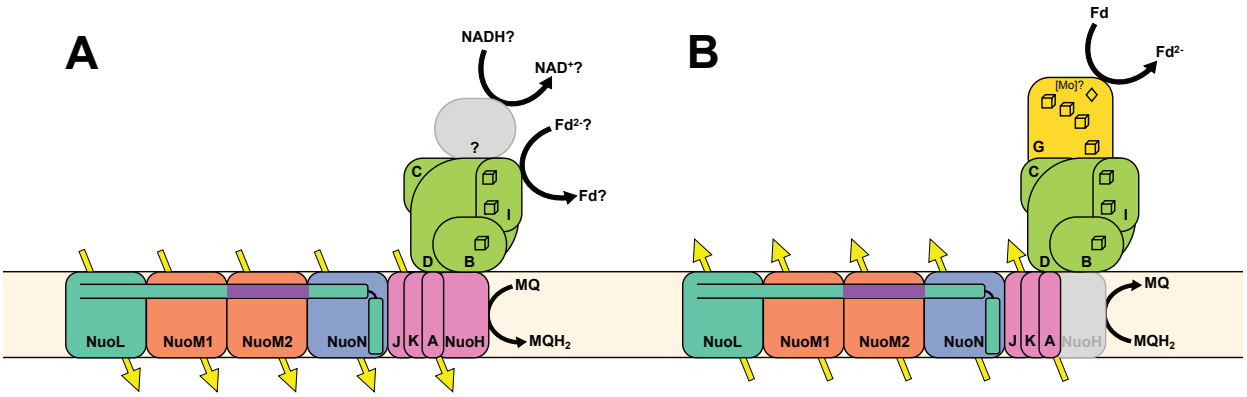


Figure 7 Possible functions of 2M complexes. (**A**) Extra proton pumping subunit may be used to extrude a fifth proton per reaction in order to conserve energy under conditions of lowered transmembrane proton motive force, such as in alkaliphilic environments or potentially as an adaptation to slow growth. (**B**) The extra proton may be used for extra driving force to conduct reverse electron transport onto low potential electron acceptors like ferredoxin (Fd) for use in carbon or nitrogen fixation.

Note: figures S1-S30 are available through CaltechDATA: 10.22002/D1.1425

Figures S1-10 High-resolution phylogenetic trees showing organisms names and gene accession numbers for NuoL built using RAxML and MrBayes as described in the Methods. Branches and leaves corresponding to members of 2M complex operons have been colored by clade (1:red, 2:green, 3:blue) and labeled numerically. Figures S1-3 are constructed from sequence alignments using Clustal Omega, MAFFT and MUSCLE, respectively using the 90% rarified dataset with LG+CAT options. Figures S4-9 are constructed from Clustal Omega alignments using the 70% identity rarified dataset using Dayhoff+CAT, Dayhoff+GAMMA, LG+CAT, LG+GAMMA, WAG+CAT, and WAG+GAMMA, respectively. Figure S10 is a tree constructed using MrBayes with an 80% identity rarified dataset with the LG substitution matrix run for 35,410,000 generations reaching a standard deviation of split frequencies of 0.022 (two runs with four chains, 1 cold, 3 heated). Figure S1 is the same tree topology used in Figure 2b. All trees show branch support with values 50% or greater.

Figures S11-20 High-resolution phylogenetic trees showing organisms names and gene accession numbers for NuoM built using RAxML and MrBayes as described in the Methods. Branches and leaves corresponding to members of 2M complex operons have been colored by clade (1:red, 2:green, 3:blue) and labeled numerically. Figures S11-13 are constructed from sequence alignments using Clustal Omega, MAFFT and MUSCLE, respectively using the 90% rarified dataset with LG+CAT options. Figures S14-19 are constructed from Clustal Omega alignments using the 70% identity rarified dataset using Dayhoff+CAT, Dayhoff+GAMMA, LG+CAT, LG+GAMMA, WAG+CAT, and WAG+GAMMA, respectively. Figure S20 is a tree constructed using MrBayes with a 70% identity rarified dataset with the LG substitution matrix run for 50,000,000 generations reaching a standard deviation of split frequencies of 0.015 (two runs with four chains, 1 cold, 3 heated). Figure S11 is the same tree topology used in Figure 3a. All trees show branch support with values 50% or greater.

Figures S21-30 High-resolution phylogenetic trees showing organisms names and gene accession numbers for NuoN built using RAxML and MrBayes as described in the Methods. Branches and leaves corresponding to members of 2M complex operons have been colored by clade (1:red, 2:green, 3:blue) and labeled numerically. Figures S21-23 are constructed from sequence alignments using Clustal Omega, MAFFT and MUSCLE, respectively using the 90% rarified dataset with LG+CAT options. Figures S24-29 are constructed from Clustal Omega alignments using the 70% identity rarified dataset using Dayhoff+CAT, Dayhoff+GAMMA, LG+CAT, LG+GAMMA, WAG+CAT, and WAG+GAMMA, respectively. Figure S30 is a tree constructed using MrBayes with a 70% identity rarified dataset with the LG substitution matrix run for 47,000,000 generations reaching a standard deviation of split frequencies of 0.025 (two runs with four chains, 1 cold, 3 heated). Figure S21 is the same tree topology used in Figure 3b. All trees show branch support with values 50% or greater.

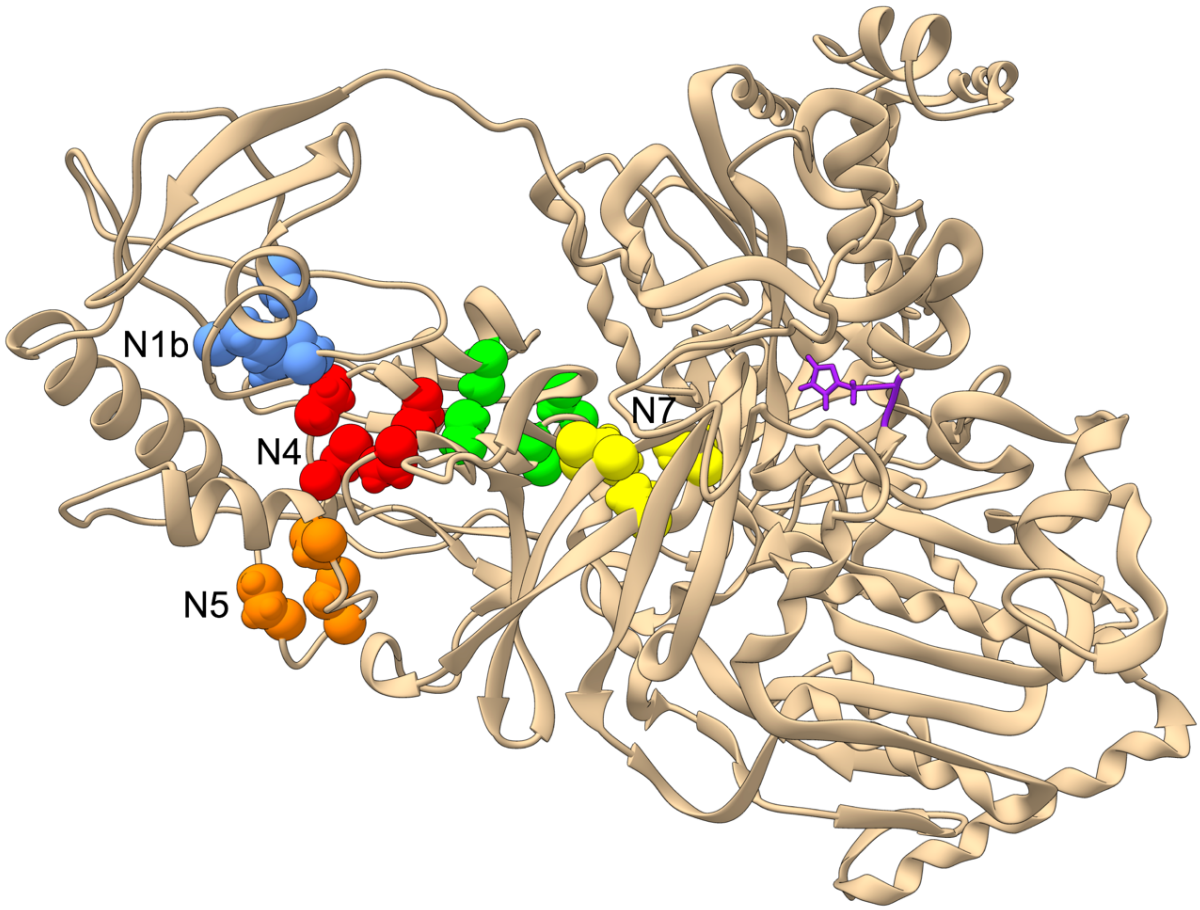
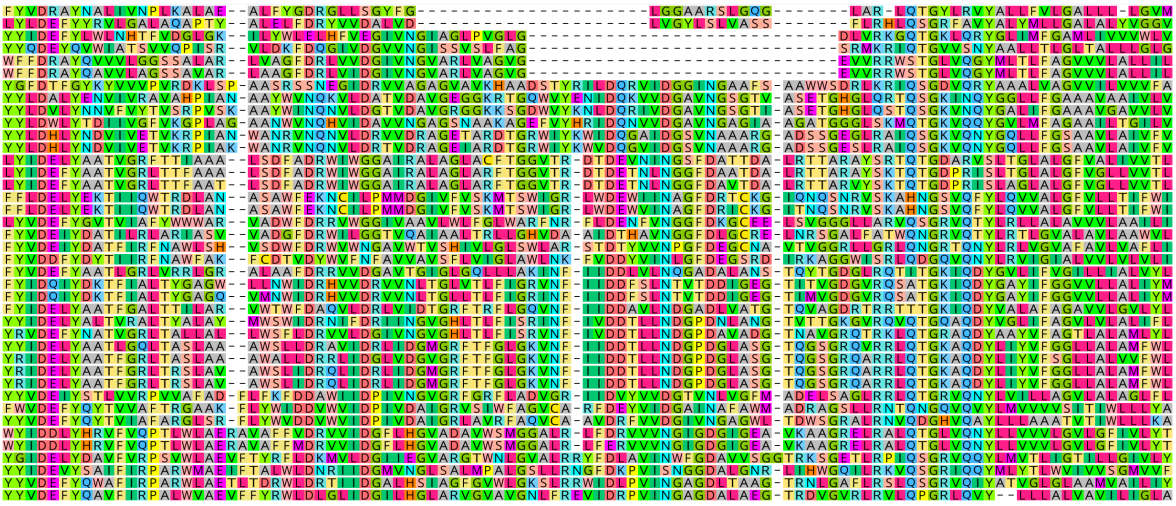


Figure S31 I-TASSER structural homology model of the NuoG protein found in the *Nitrospira defluvii* 2M complex (IMG Gene ID: 649658404) based on the formate dehydrogenase (PDB:1FDO). Cysteine residues that ligate iron sulfur clusters found in the NuoG protein of traditional complex I homologs such as Nuo in *T. thermophilus* are shown in blue, red, orange, and yellow (coordinating FeS clusters N1b, N4, N5 and N7 respectively). The extra putative FeS cluster binding cysteines are shown in green, at positions previously described in *C. jejuni* and *H. pylori*. A FeS cluster at this location would likely connect the N7 cluster with the rest of the clusters in the chain. The positions corresponding to the residues important for enzyme activity in formate dehydrogenase, SeCys140 and His141, are marked in purple. His141 is retained, but the equivalent position to SeCys140, the molybdenum atom ligand, is occupied by a glycine which cannot replace this function.

A



B

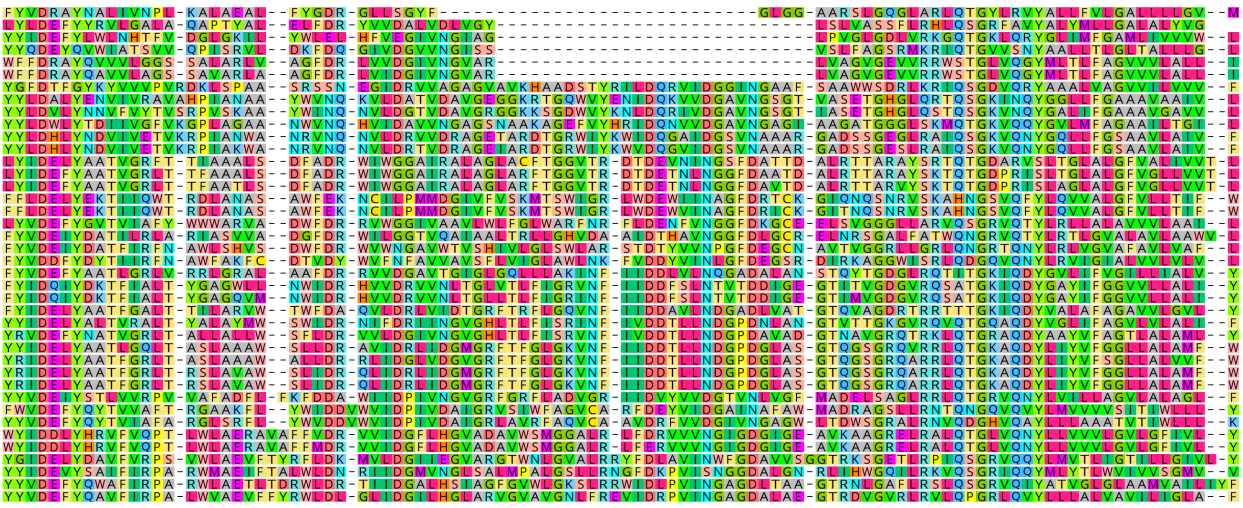


Figure S32 Sequence alignments of the NuoL gene from Clade 1 highlighting the 26-28 amino acid insertion in the amphipathic helix region. These figures correspond to the same region and gene order as shown in Figure 5a. (**A**) Clustal W alignment. (**B**) MAFFT alignment.

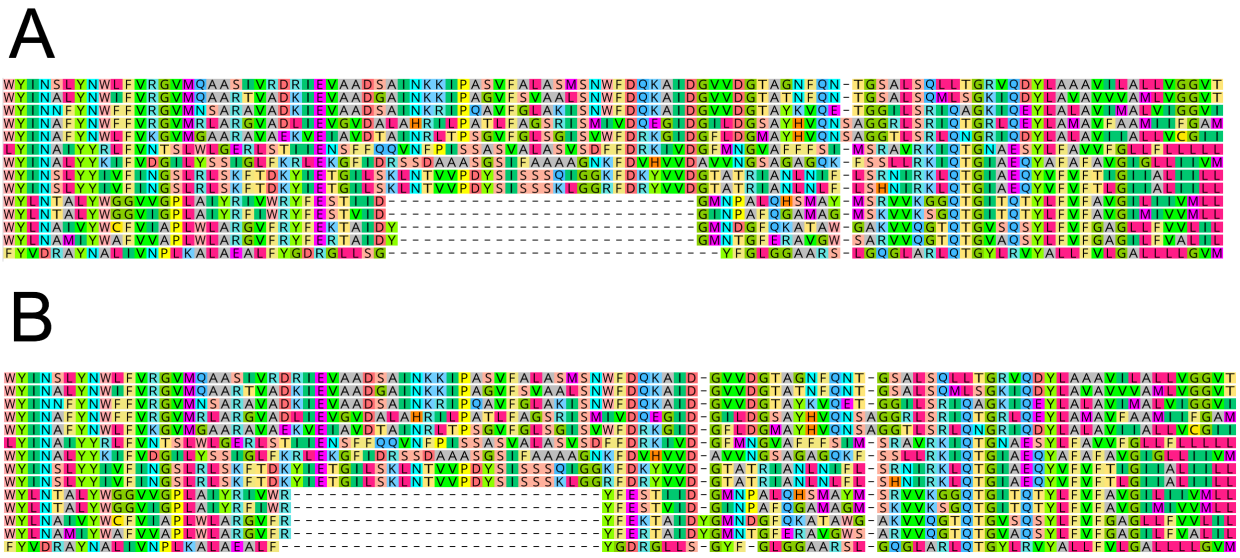


Figure S33 Sequence alignments of the NuoL gene from Clade 2 highlighting the 26-28 amino acid insertion in the amphipathic helix region. These figures correspond to the same region and gene order as shown in Figure 5b. (**A**) Clustal W alignment. (**B**) MAFFT alignment.

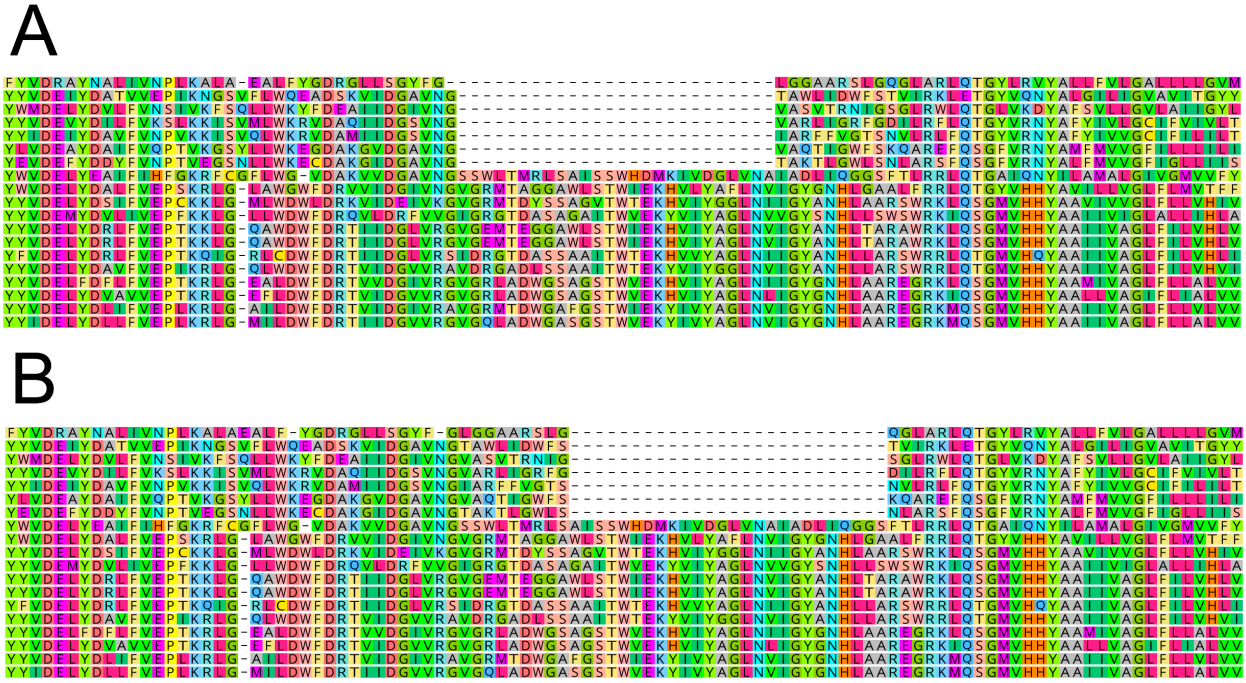


Figure S34 Sequence alignments of the NuoL gene from Clade 3 highlighting the 26-28 amino acid insertion in the amphipathic helix region. These figures correspond to the same region and gene order as shown in Figure 5c. (**A**) Clustal W alignment. (**B**) MAFFT alignment.

References

Aleem MIH. 1966. Generation of reducing power in chemosynthesis II. Energy-linked reduction of pyridine nucleotides in the chemoautotroph, Nitrosomonas europaea. *Biochimica et Biophysica Acta (BBA) - Enzymology and Biological Oxidation* **113**:216–224. doi:10.1016/S0926-6593(66)80062-0

Babauta JT, Nguyen HD, Harrington TD, Renslow R, Beyenal H. 2012. pH, redox potential and local biofilm potential microenvironments within *Geobacter sulfurreducens* biofilms and their roles in electron transfer. *Biotechnol Bioeng* **109**:2651–2662. doi:10.1002/bit.24538

Baradaran R, Berrisford JM, Minhas GS, Sazanov LA. 2013. Crystal structure of the entire respiratory complex I. *Nature* **494**:443–448. doi:10.1038/nature11871

Battchikova N, Eisenhut M, Aro E-M. 2011. Cyanobacterial NDH-1 complexes: Novel insights and remaining puzzles. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, Regulation of Electron Transport in Chloroplasts **1807**:935–944. doi:10.1016/j.bbabio.2010.10.017

Battin TJ, Kaplan LA, Denis Newbold J, Hansen CME. 2003. Contributions of microbial biofilms to ecosystem processes in stream mesocosms. *Nature* **426**:439–442. doi:10.1038/nature02152

Bäumer S, Ide T, Jacobi C, Johann A, Gottschalk G, Deppenmeier U. 2000. The F420H2 Dehydrogenase fromMethanosarcina mazei Is a Redox-driven Proton Pump Closely Related to NADH Dehydrogenases. *J Biol Chem* **275**:17968–17973. doi:10.1074/jbc.M000650200

Belevich G, Knuuti J, Verkhovsky MI, Wikström M, Verkhovskaya M. 2011. Probing the mechanistic role of the long α-helix in subunit L of respiratory Complex I from Escherichia coli by site-directed mutagenesis. *Molecular Microbiology* **82**:1086–1095. doi:10.1111/j.1365-2958.2011.07883.x

Bonanni PS, Bradley DF, Schrott GD, Busalmen JP. 2013. Limitations for current production in *Geobacter sulfurreducens* biofilms. *ChemSusChem* **6**:711–720. doi:10.1002/cssc.201200671

Bond DR, Holmes DE, Tender LM, Lovley DR. 2002. Electrode-reducing microorganisms that harvest energy from marine sediments. *Science* **295**:483–485. doi:10.1126/science.1066771

Cerqueira NMFSA, Gonzalez PJ, Fernandes PA, Moura JJG, Ramos MJ. 2015. Periplasmic Nitrate Reductase and Formate Dehydrogenase: Similar Molecular Architectures with Very Different Enzymatic Activities. *Acc Chem Res* **48**:2875–2884. doi:10.1021/acs.accounts.5b00333

Daims H, Lebedeva EV, Pjevac P, Han P, Herbold C, Albertsen M, Jehmlich N, Palatinszky M, Vierheilig J, Bulaev A, Kirkegaard RH, von Bergen M, Rattei T, Bendinger B, Nielsen PH, Wagner M. 2015. Complete nitrification by Nitrospira bacteria. *Nature* **528**:504–509. doi:10.1038/nature16461

de Beer D, Stoodley P, Roe F, Lewandowski Z. 1994. Effects of biofilm structures on oxygen distribution and mass transport. *Biotechnol Bioeng* **43**:1131–1138. doi:10.1002/bit.260431118

Deppenmeier U, Blaut M, Mahlmann A, Gottschalk G. 1990. Reduced coenzyme F420: heterodisulfide oxidoreductase, a proton- translocating redox system in methanogenic bacteria. *PNAS* **87**:9449–9453. doi:10.1073/pnas.87.23.9449

Deschamps P, Zivanovic Y, Moreira D, Rodriguez-Valera F, López-García P. 2014. Pangenome Evidence for Extensive Interdomain Horizontal Transfer Affecting Lineage Core and Shell Genes in Uncultured Planktonic Thaumarchaeota and Euryarchaeota. *Genome Biol Evol* **6**:1549–1563. doi:10.1093/gbe/evu127

Domozych DS, Domozych CR. 2008. Desmids and biofilms of freshwater wetlands: development and microarchitecture. *Microb Ecol* **55**:81–93. doi:10.1007/s00248-007-9253-y

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**:1792–1797. doi:10.1093/nar/gkh340

E. Franks A, P. Nevin K, Jia H, Izallalen M, L. Woodard T, R. Lovley D. 2009. Novel strategy for three-dimensional real-time imaging of microbial fuel cell communities: monitoring the inhibitory effects of proton accumulation within the anode biofilm. *Energy & Environmental Science* **2**:113–119. doi:10.1039/B816445B

Efremov RG, Baradaran R, Sazanov LA. 2010. The architecture of respiratory complex I. *Nature* **465**:441–445. doi:10.1038/nature09066

Efremov RG, Sazanov LA. 2012. The coupling mechanism of respiratory complex I — A structural and evolutionary perspective. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 17th European Bioenergetics Conference **1817**:1785–1795. doi:10.1016/j.bbabio.2012.02.015

Ehrich S, Behrens D, Lebedeva E, Ludwig W, Bock E. 1995. A new obligately chemolithoautotrophic, nitrite-oxidizing bacterium,Nitrospira moscoviensis sp. nov. and its phylogenetic relationship. *Arch Microbiol* **164**:16–23. doi:10.1007/BF02568729

Ehrlich HL. 2016. Uppermost lithosphere as a microbial habitatEhrlich's Geomicrobiology. Boca Raton, FL: CRC Press. pp. 55–68.

Fischer WW, Hemp J, Johnson JE. 2016. Evolution of Oxygenic Photosynthesis. *Annual Review of Earth and Planetary Sciences* **44**:647–683. doi:10.1146/annurev-earth-060313-054810

Flemming H-C, Wingender J, Szewzyk U, Steinberg P, Rice SA, Kjelleberg S. 2016. Biofilms: an emergent form of bacterial life. *Nat Rev Micro* **14**:563–575. doi:10.1038/nrmicro.2016.94

Franks AE, Glaven RH, Lovley DR. 2012a. Real-time spatial gene expression analysis within current-producing biofilms. *ChemSusChem* **5**:1092–1098. doi:10.1002/cssc.201100714

Franks AE, Glaven RH, Lovley DR. 2012b. Real-Time Spatial Gene Expression Analysis within Current-Producing Biofilms. *ChemSusChem* **5**:1092–1098. doi:10.1002/cssc.201100714

Franks AE, Nevin KP, Glaven RH, Lovley DR. 2010. Microtoming coupled to microarray analysis to evaluate the spatial metabolic status of *Geobacter sulfurreducens* biofilms. *The ISME Journal* **4**:509. doi:10.1038/ismej.2009.137

Friedrich T. 1998. The NADH:ubiquinone oxidoreductase (complex I) from Escherichia coli. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1364**:134–146. doi:10.1016/S0005-2728(98)00024-3

Friedrich T, Scheide D. 2000. The respiratory complex I of bacteria, archaea and eukarya and its module common with membrane-bound multisubunit hydrogenases. *FEBS Letters* **479**:1–5. doi:10.1016/S0014-5793(00)01867-6

Friedrich T, Steinmüller K, Weiss H. 1995. The proton-pumping respiratory complex I of bacteria and mitochondria and its homologue in chloroplasts. *FEBS Letters* **367**:107–111. doi:10.1016/0014-5793(95)00548-N

Gantzer CJ, Rittmann BE, Herricks EE. 1988. Mass transport to streambed biofilms. *Water Research* **22**:709–722. doi:10.1016/0043-1354(88)90182-0

Gao F, Zhao J, Wang X, Qin S, Wei L, Ma W. 2016. NdhV Is a Subunit of NADPH Dehydrogenase Essential for Cyclic Electron Transport in Synechocystis sp. Strain PCC 6803. *Plant Physiology* **170**:752–760. doi:10.1104/pp.15.01430

Grimaldi S, Schoepp-Cothenet B, Ceccaldi P, Guigliarelli B, Magalon A. 2013. The prokaryotic Mo/W-bisPGD enzymes family: A catalytic workhorse in bioenergetic. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, Metals in Bioenergetics and Biomimetics Systems **1827**:1048–1085. doi:10.1016/j.bbabio.2013.01.011

Hallberg ZF, Chan CH, Wright TA, Kranzusch PJ, Doxzen KW, Park JJ, Bond DR, Hammond MC. n.d. Structure and mechanism of a Hypr GGDEF enzyme that activates cGAMP signaling to control extracellular metal respiration. *eLife* **8**. doi:10.7554/eLife.43959

Hicks DB, Liu J, Fujisawa M, Krulwich TA. 2010. F1F0-ATP synthases of alkaliphilic bacteria: Lessons from their adaptations. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1797**:1362–1377. doi:10.1016/j.bbabio.2010.02.028

Holmes DE, Bond DR, O'Neil RA, Reimers CE, Tender LR, Lovley DR. 2004. Microbial communities associated with electrodes harvesting electricity from a variety of aquatic sediments. *Microb Ecol* **48**:178–190. doi:10.1007/s00248-003-0004-4

Hunte C, Zickermann V, Brandt U. 2010. Functional Modules and Structural Basis of Conformational Coupling in Mitochondrial Complex I. *Science* **329**:448–451. doi:10.1126/science.1191046

Inoue K, Leang C, Franks AE, Woodard TL, Nevin KP, Lovley DR. 2011. Specific localization of the *c*-type cytochrome OmcZ at the anode surface in current-producing biofilms of *Geobacter sulfurreducens*. *Environ Microbiol Rep* **3**:211–217. doi:10.1111/j.1758-2229.2010.00210.x

Islam FS, Gault AG, Boothman C, Polya DA, Charnock JM, Chatterjee D, Lloyd JR. 2004. Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* **430**:68–71. doi:10.1038/nature02638

Jain A, Gazzola G, Panzera A, Zanoni M, Marsili E. 2011. Visible spectroelectrochemical characterization of *Geobacter sulfurreducens* biofilms on optically transparent indium tin oxide electrode. *Electrochimica Acta*, Selected Papers from the 61st ISE Meeting, Nice, France, 2010 **56**:10776–10785. doi:10.1016/j.electacta.2011.02.073

Jiménez Otero F, Chan CH, Bond DR. 2018. Identification of different putative outer membrane electron conduits necessary for Fe(III) citrate, Fe(III) oxide, Mn(IV) oxide, or electrode reduction by *Geobacter sulfurreducens*. *Journal of Bacteriology* **200**:e00347-18. doi:10.1128/JB.00347-18

Jones AJY, Blaza JN, Varghese F, Hirst J. 2017. Respiratory Complex I in Bos taurus and Paracoccus denitrificans Pumps Four Protons across the Membrane for Every NADH Oxidized. *J Biol Chem* **292**:4987–4995. doi:10.1074/jbc.M116.771899

Jung S, Regan JM. 2007. Comparison of anode bacterial communities and performance in microbial fuel cells with different electron donors. *Appl Microbiol Biotechnol* **77**:393–402. doi:10.1007/s00253-007-1162-y

Kato Marcus A, Torres CI, Rittmann BE. 2007. Conduction-based modeling of the biofilm anode of a microbial fuel cell. *Biotechnol Bioeng* **98**:1171–1182. doi:10.1002/bit.21533

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**:3059–3066. doi:10.1093/nar/gkf436

Kirchhoff C, Cypionka H. 2017. Propidium ion enters viable cells with high membrane potential during live-dead staining. *Journal of Microbiological Methods* **142**:79–82. doi:10.1016/j.mimet.2017.09.011

Koch H, Lücker S, Albertsen M, Kitzinger K, Herbold C, Spieck E, Nielsen PH, Wagner M, Daims H. 2015. Expanded metabolic versatility of ubiquitous nitrite-oxidizing bacteria from the genus Nitrospira. *PNAS* **112**:11371–11376. doi:10.1073/pnas.1506533112

Könneke M, Bernhard AE, de la Torre JR, Walker CB, Waterbury JB, Stahl DA. 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**:543–546. doi:10.1038/nature03911

Könneke M, Schubert DM, Brown PC, Hügler M, Standfest S, Schwander T, Borzyskowski LS von, Erb TJ, Stahl DA, Berg IA. 2014. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO2 fixation. *PNAS* **111**:8239–8244. doi:10.1073/pnas.1402028111

Kopf SH, McGlynn SE, Green-Saxena A, Guan Y, Newman DK, Orphan VJ. 2015. Heavy water and 15N labelling with NanoSIMS analysis reveals growth rate-dependent metabolic heterogeneity in chemostats. *Environmental Microbiology* **17**:2542–2556. doi:10.1111/1462-2920.12752

Kopf SH, Sessions AL, Cowley ES, Reyes C, Van Sambeek L, Hu Y, Orphan VJ, Kato R, Newman DK. 2016. Trace incorporation of heavy water reveals slow and heterogeneous pathogen growth rates in cystic fibrosis sputum. *Proc Natl Acad Sci USA* **113**:E110-116. doi:10.1073/pnas.1512057112

Lebedev N, Strycharz-Glaven SM, Tender LM. 2014. Spatially resolved confocal resonant Raman microscopic analysis of anode-grown *Geobacter sulfurreducens* biofilms. *ChemPhysChem* **15**:320–327. doi:10.1002/cphc.201300984

Lee H-S, Parameswaran P, Kato-Marcus A, Torres CI, Rittmann BE. 2008. Evaluation of energy-conversion efficiencies in microbial fuel cells (MFCs) utilizing fermentable and non-fermentable substrates. *Water Res* **42**:1501–1510. doi:10.1016/j.watres.2007.10.036

Lemaire R, Webb RI, Yuan Z. 2008. Micro-scale observations of the structure of aerobic microbial granules used for the treatment of nutrient-rich industrial wastewater. *ISME J* **2**:528–541. doi:10.1038/ismej.2008.12

Levar CE, Chan CH, Mehta-Kolte MG, Bond DR. 2014. An inner membrane cytochrome required only for reduction of high redox potential extracellular electron acceptors. *mBio* **5**:e02034-14. doi:10.1128/mBio.02034-14

Levar CE, Hoffman CL, Dunshee AJ, Toner BM, Bond DR. 2017. Redox potential as a master variable controlling pathways of metal reduction by *Geobacter sulfurreducens*. *The ISME Journal* **11**:741. doi:10.1038/ismej.2016.146

Liu Y, Kim H, Franklin RR, Bond DR. 2011. Linking spectral and electrochemical analysis to monitor *c*-type cytochrome redox status in living *Geobacter sulfurreducens* biofilms. *ChemPhysChem* **12**:2235–2241. doi:10.1002/cphc.201100246

Lücker S, Nowka B, Rattei T, Spieck E, Daims H. 2013. The Genome of Nitrospina gracilis Illuminates the Metabolism and Evolution of the Major Marine Nitrite Oxidizer. *Front Microbiol* **4**. doi:10.3389/fmicb.2013.00027

Lücker S, Wagner M, Maixner F, Pelletier E, Koch H, Vacherie B, Rattei T, Damsté JSS, Spieck E, Paslier DL, Daims H. 2010. A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria. *PNAS* **107**:13479–13484. doi:10.1073/pnas.1003860107

Markowitz VM, Chen I-MA, Chu K, Szeto E, Palaniappan K, Jacob B, Ratner A, Liolios K, Pagani I, Huntemann M, Mavromatis K, Ivanova NN, Kyrpides NC. 2012. IMG/M-HMP: A Metagenome Comparative Analysis System for the Human Microbiome Project. *PLOS ONE* **7**:e40151. doi:10.1371/journal.pone.0040151

Marsili E, Rollefson JB, Baron DB, Hozalski RM, Bond DR. 2008a. Microbial biofilm voltammetry: direct electrochemical characterization of catalytic electrode-attached biofilms. *Appl Environ Microbiol* **74**:7329–7337. doi:10.1128/AEM.00177-08

Marsili E, Rollefson JB, Baron DB, Hozalski RM, Bond DR. 2008b. Microbial Biofilm Voltammetry: Direct Electrochemical Characterization of Catalytic Electrode-Attached Biofilms. *Appl Environ Microbiol* **74**:7329–7337. doi:10.1128/AEM.00177-08

Marsili E, Sun J, Bond DR. 2010. Voltammetry and growth physiology of *Geobacter sulfurreducens* biofilms as a function of growth stage and imposed electrode potential. *Electroanalysis* **22**:865–874. doi:10.1002/elan.200800007

McGlynn SE, Chadwick GL, Kempes CP, Orphan VJ. 2015. Single cell activity reveals direct electron transfer in methanotrophic consortia. *Nature* **526**:531–535. doi:10.1038/nature15512

Meier T, Morgner N, Matthies D, Pogoryelov D, Keis S, Cook GM, Dimroth P, Brutschy B. 2007. A tridecameric c ring of the adenosine triphosphate (ATP) synthase from the thermoalkaliphilic Bacillus sp. strain TA2.A1 facilitates ATP synthesis at low electrochemical proton potential. *Molecular Microbiology* **65**:1181–1192. doi:10.1111/j.1365-2958.2007.05857.x

Mitchell P. 1966. Chemiosmotic Coupling in Oxidative and Photosynthetic Phosphorylation. *Biological Reviews* **41**:445–501. doi:10.1111/j.1469-185X.1966.tb01501.x

Moparthi VK, Hägerhäll C. 2011. The Evolution of Respiratory Chain Complex I from a Smaller Last Common Ancestor Consisting of 11 Protein Subunits. *J Mol Evol* **72**:484–497. doi:10.1007/s00239-011-9447-2

Moparthi VK, Kumar B, Al-Eryani Y, Sperling E, Górecki K, Drakenberg T, Hägerhäll C. 2014. Functional role of the MrpA- and MrpD-homologous protein subunits in enzyme complexes evolutionary related to respiratory chain complex I. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1837**:178–185. doi:10.1016/j.bbabio.2013.09.012

Nevin KP, Richter H, Covalla SF, Johnson JP, Woodard TL, Orloff AL, Jia H, Zhang M, Lovley DR. 2008. Power output and columbic efficiencies from biofilms of *Geobacter sulfurreducens* comparable to mixed community microbial fuel cells. *Environ Microbiol* **10**:2505–2514. doi:10.1111/j.1462-2920.2008.01675.x

Niftrik V, A L, Fuerst JA, Damsté JSS, Kuenen JG, Jetten MSM, Strous M. 2004. The anammoxosome: an intracytoplasmic compartment in anammox bacteria. *FEMS Microbiol Lett* **233**:7–13. doi:10.1016/j.femsle.2004.01.044

Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, Mclean JS, Lasken R, Tesler G,

Alekseyev MA, Pevzner PA. 2013. Assembling Single-Cell Genomes and Mini-Metagenomes From Chimeric MDA Products. *Journal of Computational Biology* **20**:714–737. doi:10.1089/cmb.2013.0084

Polerecky L, Adam B, Milucka J, Musat N, Vagner T, Kuypers MMM. 2012. Look@NanoSIMS--a tool for the analysis of nanoSIMS data in environmental microbiology. *Environ Microbiol* **14**:1009–1023. doi:10.1111/j.1462-2920.2011.02681.x

Preiss L, Klyszejko AL, Hicks DB, Liu J, Fackelmayer OJ, Yildiz Ö, Krulwich TA, Meier T. 2013. The c-ring stoichiometry of ATP synthase is adapted to cell physiological requirements of alkaliphilic Bacillus pseudofirmus OF4. *PNAS* **110**:7874–7879. doi:10.1073/pnas.1303333110

Preiss L, Langer JD, Hicks DB, Liu J, Yildiz Ö, Krulwich TA, Meier T. 2014. The c-ring ion binding site of the ATP synthase from Bacillus pseudofirmus OF4 is adapted to alkaliphilic lifestyle. *Molecular Microbiology* **92**:973–984. doi:10.1111/mmi.12605

Ramsing NB, Kühl M, Jørgensen BB. 1993. Distribution of sulfate-reducing bacteria, O_2, and H_2S in photosynthetic biofilms determined by oligonucleotide probes and microelectrodes. *Appl Environ Microbiol* **59**:3840–3849.

Reguera G, Nevin KP, Nicoll JS, Covalla SF, Woodard TL, Lovley DR. 2006. Biofilm and nanowire production leads to increased current in *Geobacter sulfurreducens* fuel cells. *Appl Environ Microbiol* **72**:7345–7348. doi:10.1128/AEM.01444-06

Renslow RS, Babauta JT, Dohnalkova AC, Boyanov MI, Kemner KM, Majors PD, Fredrickson JK, Beyenal H. 2013. Metabolic spatial variability in electrode-respiring Geobacter sulfurreducens biofilms. *Energy Environ Sci* **6**:1827–1836. doi:10.1039/C3EE40203G

Robuschi L, Tomba JP, Busalmen JP. 2017. Proving *Geobacter* biofilm connectivity with confocal Raman microscopy. *Journal of Electroanalytical Chemistry*, SI:Professor Antonio Aldaz **793**:99–103. doi:10.1016/j.jelechem.2016.11.005

Robuschi L, Tomba JP, Schrott GD, Bonanni PS, Desimone PM, Busalmen JP. 2013. Spectroscopic slicing to reveal internal redox gradients in electricity-producing biofilms. *Angew Chem Int Ed* **52**:925–928. doi:10.1002/anie.201205440

Rollefson JB, Stephen CS, Tien M, Bond DR. 2011. Identification of an extracellular polysaccharide network essential for cytochrome anchoring and biofilm formation in *Geobacter sulfurreducens*. *J Bacteriol* **193**:1023–1033. doi:10.1128/JB.01092-10

Roy A, Kucukural A, Zhang Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protocols* **5**:725–738. doi:10.1038/nprot.2010.5

Rupprecht A, Sokolenko EA, Beck V, Ninnemann O, Jaburek M, Trimbuch T, Klishin SS, Jezek P, Skulachev VP, Pohl EE. 2010. Role of the Transmembrane Potential in the Membrane Proton Leak. *Biophysical Journal* **98**:1503–1511. doi:10.1016/j.bpj.2009.12.4301

Russell JB, Cook GM. 1995. Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiol Mol Biol Rev* **59**:48–62.

Sauer K, Camper AK, Ehrlich GD, Costerton JW, Davies DG. 2002. *Pseudomonas aeruginosa* displays multiple phenotypes during development as a biofilm. *J Bacteriol* **184**:1140–1154.

Sazanov LA. 2015. A giant molecular proton pump: structure and mechanism of respiratory complex I. *Nature Reviews Molecular Cell Biology* **16**:nrm3997. doi:10.1038/nrm3997

Schmitz O, Bothe H. 1996. The diaphorase subunit HoxU of the bidirectional hydrogenase as electron transferring protein in cyanobacterial respiration? *Naturwissenschaften* **83**:525–527. doi:10.1007/BF01141957

Schrott GD, Ordoñez MV, Robuschi L, Busalmen JP. 2014. Physiological stratification in electricity-producing biofilms of *Geobacter sulfurreducens*. *ChemSusChem* **7**:598–603. doi:10.1002/cssc.201300605

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology* **7**:539. doi:10.1038/msb.2011.75

Simon J, Klotz MG. 2013. Diversity and evolution of bioenergetic systems involved in microbial nitrogen compound transformations. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, The evolutionary aspects of bioenergetic systems **1827**:114–135. doi:10.1016/j.bbabio.2012.07.005

Sleutels THJA, Ter Heijne A, Buisman CJN, Hamelers HVM. 2012. Bioelectrochemical systems: an outlook for practical applications. *ChemSusChem* **5**:1012–1019. doi:10.1002/cssc.201100732

Snider RM, Strycharz-Glaven SM, Tsoi SD, Erickson JS, Tender LM. 2012. Long-range electron transport in *Geobacter sulfurreducens* biofilms is redox gradient-driven. *PNAS* **109**:15467–15472. doi:10.1073/pnas.1209829109

Sorokin DY, Tourova TP, Kolganova TV, Detkova EN, Galinski EA, Muyzer G. 2011. Culturable diversity of lithotrophic haloalkaliphilic sulfate-reducing bacteria in soda lakes and the description of Desulfonatronumthioautotrophicum sp. nov., Desulfonatronum thiosulfatophilum sp. nov., Desulfonatronovibrio thiodismutans sp. nov., and Desulfonatronovibrio magnus sp. nov. *Extremophiles* **15**:391–401. doi:10.1007/s00792-011-0370-7

Sorokin DYu, Tourova TP, Henstra AM, Stams AJM, Galinski EA, Muyzer G. 2008. Sulfidogenesis under extremely haloalkaline conditions by Desulfonatronospira thiodismutans gen. nov., sp. nov., and Desulfonatronospira delicata sp. nov. – a novel lineage of Deltaproteobacteria from hypersaline soda lakes. *Microbiology* **154**:1444–1453. doi:10.1099/mic.0.2007/015628-0

Spang A, Poehlein A, Offre P, Zumbrägel S, Haider S, Rychlik N, Nowka B, Schmeisser C, Lebedeva EV, Rattei T, Böhm C, Schmid M, Galushko A, Hatzenpichler R, Weinmaier T, Daniel R, Schleper C, Spieck E, Streit W, Wagner M. 2012. The genome of the ammonia-oxidizing CandidatusNitrososphaera gargensis: insights into metabolic versatility and environmental adaptations. *Environ Microbiol* **14**:3122–3145. doi:10.1111/j.1462-2920.2012.02893.x

Steidl RJ, Lampa-Pastirk S, Reguera G. 2016. Mechanistic stratification in electroactive biofilms of *Geobacter sulfurreducens* mediated by pilus nanowires. *Nat Commun* **7**:12217. doi:10.1038/ncomms12217

Stephen CS, LaBelle EV, Brantley SL, Bond DR. 2014a. Abundance of the multiheme *c*-type cytochrome OmcB increases in outer biofilm layers of electrode-grown *Geobacter sulfurreducens*. *PLOS ONE* **9**:e104336. doi:10.1371/journal.pone.0104336

Stephen CS, LaBelle EV, Brantley SL, Bond DR. 2014b. Abundance of the Multiheme c-Type Cytochrome OmcB Increases in Outer Biofilm Layers of Electrode-Grown Geobacter sulfurreducens. *PLoS ONE* **9**:e104336. doi:10.1371/journal.pone.0104336

Stevenson BS, Eichorst SA, Wertz JT, Schmidt TM, Breznak JA. 2004. New Strategies for Cultivation and Detection of Previously Uncultured Microbes. *Appl Environ Microbiol* **70**:4748–4755. doi:10.1128/AEM.70.8.4748-4755.2004

Stieglmeier M, Klingl A, Alves RJE, Rittmann SK-MR, Melcher M, Leisch N, Schleper C. 2014. Nitrososphaera viennensis gen. nov., sp. nov., an aerobic and mesophilic, ammonia-oxidizing archaeon from soil and a member of the archaeal phylum Thaumarchaeota. *International Journal of Systematic and Evolutionary Microbiology* **64**:2738–2752. doi:10.1099/ijs.0.063172-0

Stoodley P, Dodds I, Boyle JD, Lappin-Scott HM. 1998. Influence of hydrodynamics and nutrients on biofilm structure. *J Appl Microbiol* **85 Suppl 1**:19S-28S. doi:10.1111/j.1365-2672.1998.tb05279.x

Stoodley P, Sauer K, Davies DG, Costerton JW. 2002. Biofilms as complex differentiated communities. *Annu Rev Microbiol* **56**:187–209. doi:10.1146/annurev.micro.56.012302.160705

Summers ZM, Fogarty HE, Leang C, Franks AE, Malvankar NS, Lovley DR. 2010. Direct exchange of electrons within aggregates of an evolved syntrophic coculture of anaerobic bacteria. *Science* **330**:1413–1415. doi:10.1126/science.1196526

Sun D, Chen J, Huang H, Liu W, Ye Y, Cheng S. 2016. The effect of biofilm thickness on electrochemical activity of *Geobacter sulfurreducens*. *International Journal of Hydrogen Energy*, Special Issue: Hydrogen and Fuel Cell Developments: A special issue on the 8th International Conference on Sustainable Energy and Environmental Protection (SEEP 2015), 11–14 August 2015, Paisley, Scotland, UK **41**:16523–16528. doi:10.1016/j.ijhydene.2016.04.163

Sun D, Cheng S, Wang A, Li F, Logan BE, Cen K. 2015. Temporal-spatial changes in viabilities and electrochemical properties of anode biofilms. *Environ Sci Technol* **49**:5227–5235. doi:10.1021/acs.est.5b00175

Sun D, Cheng S, Zhang F, Logan BE. 2017. Current density reversibly alters metabolic spatial structure of exoelectrogenic anode biofilms. *Journal of Power Sources* **356**:566–571. doi:10.1016/j.jpowsour.2016.11.115

Tadanier CJ, Schreiber ME, Roller JW. 2005. Arsenic mobilization through microbially mediated deflocculation of ferrihydrite. *Environ Sci Technol* **39**:3061–3068. doi:10.1021/es048206d

Talavera G, Castresana J, Kjer K, Page R, Sullivan J. 2007. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. *Syst Biol* **56**:564–577. doi:10.1080/10635150701472164

Teal TK, Lies DP, Wold BJ, Newman DK. 2006. Spatiometabolic stratification of *Shewanella oneidensis* biofilms. *Appl Environ Microbiol* **72**:7324–7330. doi:10.1128/AEM.01163-06

Tejedor Sanz S, Fernández Labrador P, Hart S, Torres CI, Esteve-Núñez A. 2018. *Geobacter* dominates the inner layers of a stratified biofilm on a fluidized anode during brewery wastewater treatment. *Front Microbiol* **9**:387. doi:10.3389/fmicb.2018.00378

Tender LM, Reimers CE, Stecher HA, Holmes DE, Bond DR, Lowy DA, Pilobello K, Fertig SJ, Lovley DR. 2002. Harnessing microbially generated power on the seafloor. *Nat Biotechnol* **20**:821–825. doi:10.1038/nbt716

Torres CI, Kato Marcus A, Rittmann BE. 2008. Proton transport inside the biofilm limits electrical current generation by anode-respiring bacteria. *Biotechnol Bioeng* **100**:872–881. doi:10.1002/bit.21821

Torres-Bacete J, Sinha PK, Matsuno-Yagi A, Yagi T. 2011. Structural Contribution of C-terminal Segments of NuoL (ND5) and NuoM (ND4) Subunits of Complex I from Escherichia coli. *J Biol Chem* **286**:34007–34014. doi:10.1074/jbc.M111.260968

Valentine DL. 2007. Opinion: Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nature Reviews Microbiology* **5**:nrmicro1619. doi:10.1038/nrmicro1619

Vinothkumar KR, Zhu J, Hirst J. 2014. Architecture of mammalian respiratory complex I. *Nature* **515**:80–84. doi:10.1038/nature13686

Waite DW, Vanwonterghem I, Rinke C, Parks DH, Zhang Y, Takai K, Sievert SM, Simon J, Campbell BJ, Hanson TE, Woyke T, Klotz MG, Hugenholtz P. 2017. Comparative Genomic Analysis of the Class Epsilonproteobacteria and Proposed Reclassification to Epsilonbacteraeota (phyl. nov.). *Front Microbiol* **8**. doi:10.3389/fmicb.2017.00682

Wang F, Gu Y, O'Brian JP, Yi SM, Yalcin SE, Srikanth V, Shen C, Vu D, Ing NL, Hochbaum AI, Egelman EE, Malvankar NS. 2019. Structure of microbial nanowires reveals stacked hemes that transport electrons over micrometers. *Cell* **177**:361–369.

Weerakoon DR, Olson JW. 2008. The Campylobacter jejuni NADH:Ubiquinone Oxidoreductase (Complex I) Utilizes Flavodoxin Rather than NADH. *J Bacteriol* **190**:915–925. doi:10.1128/JB.01647-07

Wegener G, Krukenberg V, Riedel D, Tegetmeyer HE, Boetius A. 2015. Intercellular wiring enables electron transfer between methanotrophic archaea and bacteria. *Nature* **526**:587–590. doi:10.1038/nature15733

Welte C, Deppenmeier U. 2014. Bioenergetics and anaerobic respiratory chains of aceticlastic methanogens. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, 18th European Bioenergetics Conference 2014 Lisbon, Portugal **1837**:1130–1147. doi:10.1016/j.bbabio.2013.12.002

Welte C, Deppenmeier U. 2011. Membrane-Bound Electron Transport in Methanosaeta thermophila. *J Bacteriol* **193**:2868–2870. doi:10.1128/JB.00162-11

Wilmes P, Remis JP, Hwang M, Auer M, Thelen MP, Banfield JF. 2009. Natural acidophilic biofilm communities reflect distinct organismal and functional organization. *The ISME Journal* **3**:266–270. doi:10.1038/ismej.2008.90

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. 2015. The I-TASSER Suite: protein structure and function prediction. *Nat Meth* **12**:7–8. doi:10.1038/nmeth.3213

Zacharoff L, Chan CH, Bond DR. 2016. Reduction of low potential electron acceptors requires the CbcL inner membrane cytochrome of *Geobacter sulfurreducens*. *Bioelectrochemistry* **107**:7–13. doi:10.1016/j.bioelechem.2015.08.003

Zhang Y. 2008. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**:40. doi:10.1186/1471-2105-9-40